Trade Reform, Oligopsony, and Labor Market Distortions: Theory and Evidence*

Hoang Pham†
Oregon State University

This Draft: March, 2021

Abstract

In a heterogeneous-firm model with oligopsonistic local labor markets, this paper shows that opening up to trade can affect distortions in such markets. These distortions arise because firms are large and able to exercise market power over their local workers. Using a panel dataset of Chinese manufacturing firms from 1998-2007, I measure firm-level labor market distortions and examine their evolution following China's trade policy reform in 2001. I find that labor market distortions are pervasive and China's trade policy reforms have led to a substantial net reduction of the distortions, with large effects working through the liberalization of input tariffs.

JEL Codes: F12, F14, F16, F61, D43, J42, L13, L22
JEL Keywords: International Trade, Labor Market, Oligopsony, Productivity

---

*I am indebted to Devashish Mitra, Mary E. Lovely and Alfonso Flores-Lagunes for their continuous guidance and support. I would also like to thank Kristy Buzard, Arindrajit Dube, Michael Gechter, Ross Jestrab, Leyla Karakas, Derek Laing, Yoonseok Lee, Michelle Liu, Philip Luck, Peter Morrow, Cong S. Pham, Abdulaziz B. Shifa, Alexandre Skiba, Gary Solon and seminar paticipants at the 2020 European Winter Meeting of the Econometric Society, Midwest International Trade Meetings, Oregon State University, University of Auckland, Syracuse University, Competition Dynamics for helpful comments and discussions. All errors are my own.

†Email: hoang.pham@oregonstate.edu

1 Introduction

The impact of international trade policy on labor market outcomes is a central topic in the international economics literature (Goldberg and Pavcnik (2016)). Although voluminous, the majority of empirical work has been based primarily on the theoretical premise that firms behave competitively in the labor market.[1] This premise stands in stark contrast to a recent empirical labor economics literature which documents that firms possess some degree of market power in the labor market, and thus, can inflict distortionary effects on the economy by engaging in non-competitive conduct therein (see for examples Card et al. (2018), Berger, Herkenhoff and Mongey (2019)).[2] Since labor market power is closely tied to a firm's performance, which is in turn affected by trade, this paper examines whether trade policy can affect competition in the labor market and thus, alter labor market outcomes through this channel.

Providing a compelling answer to this question is difficult for two reasons. First, from a theoretical point of view, it is not obvious how firms, playing the central role as mediators, transmit trade shocks in the product market to the labor market. In the presence of labor market imperfections, characterizing this transmission requires making explicit assumptions about competition structures in *both* the product and labor markets, the latter often missing in theoretical trade models. Second, from an empirical perspective, firms' distortions in the labor market are not directly observable from data, and therefore, measuring the distortions requires a consistent methodology. This paper offers a novel approach to both of these problems and provides an estimate of the impact of trade policy on the labor market power of firms, using Chinese firm-level data, with China's accession to the World Trade Organization (WTO) in 2001 as a historical policy experiment.

Formally, my analysis in this paper delivers three key contributions. I first develop a theoretical framework to formalize the notion of labor market distortion at the firm level and explain the mechanism through which trade policy affects firms' competitive behavior in the labor market. A distinct feature of this theoretical framework is that it embeds a generic oligopsony competition structure in the labor market into a workhorse trade model with heterogeneous firms as in Melitz (2003), and allows entry and exit of firms to affect competition within a local labor market. Second, guided by the theory, I propose two complementary approaches to empirically measure labor market distortion at the firm level:

[1]See for examples the canonical models of international trade with heterogeneous firms, such as those in Melitz (2003), Bernard et al. (2003), Melitz and Ottaviano (2008), and Atkeson and Burstein (2008).

[2]Earlier discussions and empirical evidence on firms' labor market power in the labor economics literature can be found in Boal and Ransom (1997), Manning (2003), Staiger, Spetz and Phibbs (2010), Ashenfelter, Farber and Ransom (2010), and Dube et al. (2020).

(1) a production function estimation approach; and (2) a regression approach that exploits a unique exogenous demand shifter in China's context, namely the US-China Trade Policy Uncertainty (TPU) shock. These two measures not only serve to cross-validate each other, but also help to quantify the magnitude of the distortion that is entirely caused by the labor market power of firms, relative to all other forms of distortion. Finally, using the resulting measures of labor market distortion, I establish a causal link between China's trade policy reform, specifically reductions in both output tariffs and input tariffs, and the consequential changes in labor market distortions.

To develop my results, I start off by deriving a "reduced-form" representation of the labor market distortion from a simple firm's profit maximization problem, following the modeling convention in the misallocation literature (Hsieh and Klenow (2009), Liu (2019)). A key result from this representation is that all sources of distortion in the labor market can be nonparametrically summarized by the ratio between the equilibrium marginal revenue product of labor ($MRPL$) and wage (w) paid by the firm, which I henceforth refer to as the *overall distortion*. Since labor market distortion can arise due to a variety of sources, I decompose the overall distortion into two main components: the *exogenous distortion* versus the *endogenous distortion*. The exogenous distortion reflects the inefficient policy features of the labor market, such as labor regulations and institutional constraints, and generally does not respond to firm-level idiosyncratic shocks. On the other hand, the endogenous distortion arises due to firm's labor market power, and thus can potentially be altered by trade shocks.

How does trade policy affect firms' labor market power distortion? To answer this central question and provide theoretical guidance for my empirical analysis, I develop a tractable trade model with heterogeneous firms and oligopsonistic competition in the labor market.[3] The model has two distinct features. First, firms are assumed to be atomistic and compete monopolistically within the national product market, as in Melitz (2003). Nonetheless, firms' locations are distributed over a continuum of local labor markets, within which firms are large employers and employ only local workers. Within each local labor market, I explicitly model the oligopsony structure using a nested constant elasticity of substitution (CES) labor supply system, building on the recent microfoundations from the labor economics literature, particularly as in Berger, Herkenhoff and Mongey (2019). Second, I allow for entry and exit of firms following a sectoral trade shock, which consequently serves as the main mechanism through which trade policy affects the distortion in the labor market. This modeling approach is motivated by two empirical patterns that I observe in the data: (1) there are massive entries and exits following trade liberalization within a local labor

[3] As will be shown later, the model shares similar flavors to trade models with oligopolistic competition in the product market, for example as in Atkeson and Burstein (2008) and Edmond, Midrigan and Xu (2015).

market, and (2) firms' local labor market share responds to trade policy.[4] To the best of my knowledge, this is the first trade model that incorporates endogenous entry and exit in an oligopsony context with heterogenous firms and allows them to affect labor market structure.

My model provides clear and intuitive predictions for the impact of trade policy. To begin with, in this model, more productive firms always have larger local labor market shares and exercise more market power over their workers. Starting from an initial equilibrium, when the Home country opens up to trade by lowering its output tariffs, the competitive pressure from Foreign imports reduces each firm's profit. Those firms at the margin, i.e. firms with productivity level near the operating threshold, reoptimize and decide to: (1) stay or (2) exit the market, whereas non-incumbent firms may decide to (3) enter. Since profit generally decreases due to competitive trade shocks, the least productive firms exit; labor market share is reallocated towards more productive firms; and thus, the average distortion increases. On the other hand, when the Home country lowers input tariffs, it reduces production cost for all firms that use foreign inputs, increases each firm's profit, and induces entry of less productive firms into the market. As these firms gain market share, the average labor market distortion decreases.

Empirically, I tackle the measurement of the labor market distortion with two approaches. My baseline approach exploits the "reduced-form" representation to measure the overall distortion. More specifically, since the distortion can be captured by the ratio between the marginal revenue product of labor ($MRPL$) and wage (w) paid by the firm, measurement of the distortion translates naturally into estimation of $MRPL$, which can typically be obtained by identifying a *revenue* production function. To this end, I adopt a production function estimation technique recently introduced in the industrial organization literature by Gandhi, Navarro and Rivers (2020) (henceforth, GNR) to estimate a general, nonparametric production function using Chinese firm-level production data. The GNR method is distinguished from other existing methods, such as Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves and Frazer (2015), in that its identification is grounded in economic theory, rather than in functional-form assumptions of the production function.

In the second approach to measurement, I exploit a unique exogenous demand shock to Chinese firms, namely the US-China Trade Policy Uncertainty (TPU) shock, to identify the endogenous distortion. This approach combines the insights from the pass-through and trade policy uncertainty literatures, as in Amiti, Itskhoki and Konings (2019) and Handley and Limão (2017), Pierce and Schott (2016), respectively.[5] The intuition for the identification

[4] Benmelech, Bergman and Kim (2020) document similar patterns for US manufacturing sectors. They find that employer concentration, measured by the Herfindahl-Hirschman Index, decreases substantially following the import competition shock from China in the early of the 2000s.

[5] Following China's accession to WTO in 2001, the US moves China *permanently* from the "Column 2"

is that the TPU shock technically acts as an exogenous labor demand shock. Therefore, by observing the relational responses of firms in terms of wage and employment, I can trace out the slope and thus, elasticity of the labor supply curve. Furthermore, I allow for the pattern of the response to be dependent on firms' local labor market share, which consequently permits a measure of share-dependent firm-level endogenous distortion computed from the regression estimates.

From the empirical estimates of the labor market distortion, two main findings emerge. First, the baseline estimates of the *overall distortion* from production function estimation indicate that labor market distortion is pervasive among the Chinese manufacturing sectors, with the average magnitude of the distortion implying a 47% pass-through rate of an idiosyncratic productivity shock to wage.[6] The *endogenous distortion*, which arises due to firms' labor market power, accounts for almost 76% of the overall distortion. Importantly, throughout my 10-year sample of Chinese firms, I also find that key moments such as the mean, median, and dispersion of the distortion decrease over time.

With the obtained measures, I empirically assess the impact of trade policy on the labor market distortion of firms, using China's accession to WTO in 2001 as a major shift in the country's trade policy regime. My empirical model compares changes in the measured distortion between firms located within the same local labor market, yet exposed differentially to trade shocks due to their industry affiliations, following the popular approach in the empirical trade literature, as in Pavcnik (2002), Amiti and Konings (2007), Topalova and Khandelwal (2011), and Brandt et al. (2017). I find strong empirical support for my theoretical predictions. Qualitatively, increased import competition due to lower output tariffs leads to an increase in labor market distortion. Even though the effect is consistent with the theory, the magnitude and the statistical significance of this effect are small. On the other hand, access to cheaper inputs due to lower input tariffs causes a significant decrease in labor market distortion. I estimate that China's lowering of input tariffs during the sample period from 1998-2007 reduced labor market distortion by 2.95% on average.

Summarizing the results, my analysis suggests that labor market distortions were substantial and pervasive in China's manufacturing sector during the sample period. An important finding is that local labor market power accounts for a major part of overall labor market distortion. Trade policy, by changing the aggregate profitability of a sector, can induce entry and exit of firms across local labor markets, and thus, affects distortion in these markets.

tariffs to the Most Favored Nation (MFN) tariffs, and thus eliminates the possibility that China might face surprisingly high "Column 2" tariffs rather than the MFN tariffs, which are already granted to China prior to its WTO accession. It is also important to note that this TPU shock of US towards China is distinct and uncorrelated to China's own trade policy, which is the main policy focus of this paper.

[6] This pass-through rate would be 100% in an environment where there is no labor market distortion.

These findings suggest novel effects of trade policy that deviate from the conventional trade models.

Related Literature Theoretically, my paper builds on the international trade and labor market imperfections literature. Most related to my modeling approach of the constant elasticity of substitution (CES) labor supply system is the study by Berger, Herkenhoff and Mongey (2019). In their paper, the CES labor supply system is micro-founded from the discrete choice model of each individual worker, much like how the CES demand system is derived.[7] My contribution is to embed this CES labor supply system into a canonical trade model of Melitz (2003), and allow trade policy to affect the local labor market competition through entry and exit of firms. By modeling the product market as monopolistic competition with constant markups, I can abstract from the complication of strategic interactions in the product market and specify a simple equilibrium selection rule to close the model, following the modeling technique in Atkeson and Burstein (2008), Eaton, Kortum and Sotelo (2012), Edmond, Midrigan and Xu (2015) and most recently Gaubert and Itskhoki (2021). There are a few previous studies that also integrate labor market imperfections into trade models with heterogeneous firms. Most recently, MacKenzie (2019) develops and estimates a quantitative trade model with oligopoly in the product market and oligopsony in the labor market, using Indian plant-level data. However, due to the complexity of the strategic interactions in both markets, he has to assume that the number of active firms in the market is exogenous, and trade affects the labor market power through changes in product market power, a mechanism distinct from my model. In another closely related study, Heiland and Kohler (2019) examines a theoretical model with oligopoly and oligopsony, and allows for endogenous exits due to trade. In their framework, firms are homogenous and oligopsony arises due to horizontal worker heterogeneity. Ranjan and Rodriguez-Lopez (2019) specify a trade model with monopolistic competition in the product market and monopsonistic competition in the labor market to reexamine the welfare implications of trade. Helpman, Itskhoki and Redding (2010) and Amiti and Davis (2011) also incorporate labor market imperfections into trade models. Nonetheless, a common feature of these studies is that, because labor market distortion is captured by fixed parameters, there is little room for trade to endogenously affect the distortion.

Methodologically, my paper is related to the productivity, markup, and pass-through estimation literatures. Recently, productivity estimation has been used to measure and study

[7] In a related framework, Card et al. (2018) also develops a microfoundation for the monopsonistic competition structure of the labor market based the discrete choice framework. However, the monopsonistic competition is not well-suited for my study because by setup, the labor market distortion is assumed to be constant and cannot be affected by trade.

market power in product market, for example, as in De Loecker and Warzynski (2012), Flynn, Gandhi and Traina (2019). In trade literature, the impact of trade policy on market power in product market has attracted a large attention, including De Loecker (2011), De Loecker et al. (2016), Brandt et al. (2017). A growing literature has also adopted the productivity framework to study market power of firms in factor markets, including Morlacco (2020), Brooks et al. (2021), Dobbelaere and Wiersma (2019).[8] A common estimation framework used in these studies is the method developed by a series of papers, including Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves and Frazer (2015) (henceforth, ACF). A drawback to the ACF approach is that identification of the production function requires a Leontief functional-form assumption, which in the Chinese firm-level context significantly overestimates the labor-ouput elasticity and produces unrealistically large measures of the labor market distortion.[9]

In this paper, I adopt the Gandhi, Navarro and Rivers (2020)'s method (henceforth, GNR) to consistently estimate labor market distortion at the firm level. My estimation procedure does not impose functional-form assumptions and produces more reasonable estimates.[10] To complement the production function approach, I also adopt insights from the pass-through estimation literature, as in Amiti, Itskhoki and Konings (2019), to measure the endogenous component of the distortion. In this literature, the pass-through of international shocks to firm-level domestic prices is allowed to be dependent on firm's market share within an industry. The local labor market share plays a similar role in my analysis and permits a variable pass-through rate from a productivity shock to wage paid by a firm.

Finally, my analysis of the labor market distortion in this paper is related to a large literature on resource misallocation due to market imperfections, including Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Edmond, Midrigan and Xu (2015), and Liu (2019), among others. In my analysis, labor market distortion is a form of "labor tax" that impedes a more efficient reallocation of labor in response to trade shocks. My empirical results suggest that labor market distortion is large and on par with product market distortion, yet only the latter of which has been incorporated into the welfare calculations of trade, as in Arkolakis, Costinot and Rodriguez-Clare (2012), Arkolakis et al. (2018), and others.

[8] Dobbelaere and Wiersma (2019) is a contemporaneous study that empirically investigates the impacts of trade policy on distortions in both product and labor market, exploiting the same empirical context as in this paper. The difference between this paper and their paper is in terms of theory and method. The analysis in this paper provides theory guidance and uses fundamentally different approaches to both measure distortions as well as identify the impacts of trade policy.

[9] I estimate labor market distortion using the ACF approach and show these results in Table D2 in the Appendix.

[10] Recently, Lu, Sugita and Zhu (2019) also adopt the GNR approach to jointly estimate a constant markup and variable wage markdowns using Chinese firm-level data. They use these measures study the impact of foreign investment on labor market distortions using a shift-share empirical approach.

2 Labor Market Distortion and Theoretical Motivations

This section formalizes the notion of labor market distortion and develops a tractable model to study the impact of a trade policy reform on firms' competitive conduct in the labor market. To achieve these goals, in section 2.1, I derive a "reduced-form" representation of the labor market distortion from firms' profit maximization problem, and conceptually distinguish between the exogenous versus the endogenous components of the distortion. This "reduced-form" representation is important, because it provides me with a nonparametric framework to measure the *overall distortion* from production data, without having to impose any structure on either the product or the labor market. In section 2.2, I develop a richer model to capture the main sources that give rise to the endogenous distortion of firms. The model also serves to clarify the mechanism through which firms endogenously respond to changes in the competitive environment due to trade shocks.

2.1 Labor Market Distortion

Labor market distortion reflects inefficiencies in the labor market. These inefficiencies can arise from policy interventions and institutional constraints, or via externalities and non-competitive conducts of firms in the labor market. Regardless of the source and the interpretation, labor market distortion can be viewed as a form of "labor tax" and has two important properties. First, it raises the marginal wage cost: for every dollar spent on an additional worker, firms have to pay an extra amount to cover the distortion payment. Second, this extra payment reflects the size of inefficiency that inflicts a dead weight loss on the economy due to firms' suboptimal level of production.

To be as general as possible about its formulation, I introduce labor market distortion into firm i's profit maximization problem by classifying the distortion into two major components: (1) an exogenous policy distortion and (2) an endogenous labor market power distortion (henceforth, the *exogenous distortion* and the *endogenous distortion* repsectively).[11] The policy distortion is captured by a distortionary wedge χ^x and manifests as a uniform "labor tax" imposed on the labor supply curve facing *all* firms within the same labor market. This approach follows the modeling convention in the misallocation literature, for example as in Hsieh and Klenow (2009), Liu (2019). On the other hand, the labor market power distor-

[11] The terms "exogenous" and "endogenous" here are used to describe the fact that the former type of distortion generally does not vary across firms and does not respond to firm-level idiosyncratic shocks. On the other hand, the latter type of distortion is dependent on firms' optimizing decision, and thus generally responds to firm-level shocks.

tion arises due to an upward-sloping labor supply curve facing each individual firm $w(L_i)$. A classic example of this distortion is when the firm faces a monopsonistic labor market (Manning (2003)). In such an environment, the wage offer becomes a nontrivial primitive function of labor supply and thus, gives the firm some degree of freedom to set the wage. In what follows, I define the firm's problem, and derive a "reduced-form" representation to summarize all the distortions in the labor market for each firm.

Firm's Problem Firm i maximizes its profit by solving the following problem:

$$\max_{L_i} \Pi(L_i) = R(L_i) - (1 + \chi^x)w(L_i)L_i, \tag{1}$$

where $R(L_i)$ is the revenue of firm i, as a function of labor factor L_i. $w(L_i)$ is an arbitrary labor supply function facing firm i. χ^x represents the policy distortion common to all firms. Notice that this setup is general, in the sense that it does *not* assume that the labor market is imperfect, nor that workers are homogenous across firms. Such generality is preserved in this simple framework because the labor supply function $w(L_i)$ is allowed to be a constant, i.e. $w(L_i) \equiv w \ \forall L_i$, which corresponds to the case of perfect competition. On the other hand, firms with heterogeneous workforce end up solving the same problem as in (1) if workers' average ability acts as a Hicks-neutral productivity shock for each firm, as in Helpman, Itskhoki and Redding (2010). For the most part of this paper, however, I abstract from this issue, motivated by the fact that the majority of China's manufacturing employment are low-skilled.[12]

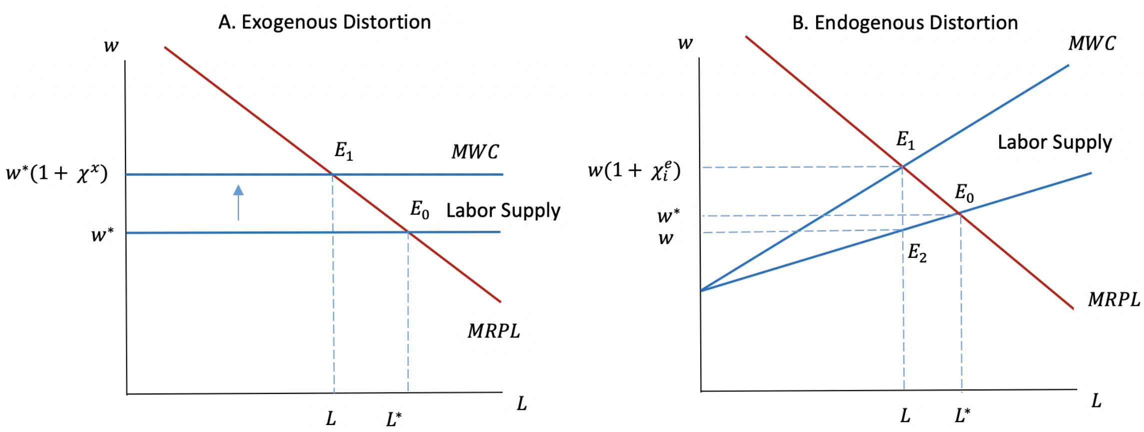
First-order condition (FOC) of the problem in (1) yields the following expression:

$$\begin{aligned} MRPL_i &= (1 + \chi^x)(1 + \underbrace{\frac{\partial w_i}{\partial L}\frac{L}{w_i}})w_i \\ &= (1 + \chi^x) \ \ (\ 1 \ + \ \ \chi_i^e \ )w_i, \end{aligned} \tag{2}$$

where $MRPL_i \equiv \frac{\partial R_i(.)}{\partial L}$ and w_i are respectively the marginal revenue product of labor and the wage paid by firm i in equilibrium. Let us further denote $\chi_i^e \equiv \frac{\partial w_i}{\partial L}\frac{L}{w_i}$ as the inverse elasticity of labor supply curve. From the expression in (2), it is clear that distortions generate a wedge between the equilibrium marginal revenue product of labor ($MRPL_i$) and wage (w_i) paid by the firm. In a distortion-free economy where there is no policy distortion and the labor market is perfectly competitive, i.e. $\chi^x = 0$ and $\frac{\partial w_i}{\partial L} = 0$, $MRPL_i$ is set to equalize

[12] In 2004, 88.4% of Chinese manufacturing workers do not have a college degree, of which 53% have only secondary degree, and 35.4% finish high school. Source: author's calculations using Chinese firm-level data in 2004.

Figure 1: Exogenous versus Endogenous Distortion in the Labor Market



Note: The figure illustrates the exogenous versus the endogenous features of labor market distortions. Panel A (left) shows the effect of the exogenous distortion that acts as a uniform "labor tax" (χ^x) on all firms within the same local labor market. Panel B (right) shows the effect of the endogenous distortion (χ_i^e) that varies with firm's size.

the wage w_i as the firm seeks to maximize profit. The roles of the exogenous distortion χ^x and the endogenous distortion χ_i^e in firm's problem are illustrated in the panels A and B of Figure 1.

Let us further denote $\tilde{\chi}_i \equiv (1+\chi^x)(1+\chi_i^e)$. From (2), $\tilde{\chi}_i$ summarizes all the distortions in the labor market, which I henceforth refer to as the *overall distortion*, and has a "reduced-form" representation as:

$$\tilde{\chi}_i = \frac{MRPL_i}{w_i}. \tag{3}$$

The expression in equation (3) is crucial for this study for two reasons. First, the magnitude of $\tilde{\chi}_i$ directly indicates the pass-through rate of an idiosyncratic labor demand shock to the firm's wage, and hence, carries information about the response of the wage distribution across firms to external shocks. Second, equation (3) provides a straight-forward guidance to estimate the overall distortion based on production data: given the data on wage, the estimation problem of $\tilde{\chi}_i$ transalates naturally to the estimation problem of $MRPL_i$, which I accomplish by identifying a revenue production function. This procedure is nonparametric, in the sense that I do not have to impose any assumptions on the product market, the labor market or production technology.[13]

It is important to point out that in the environment set up by firm's problem in (1)

[13] In the labor economics literature, $\tilde{\chi}_i$ is usually refered to as either the degree of exploitation of workers by firms (Pigou (1924), Robinson (1969), Boal and Ransom (1997)) or the inverse of the wage markdown.

combined with the FOC expression in (2), the only source of endogenous wage variation in this framework comes from the non-horizontal labor supply curve. In addition, because the elasticity of labor supply depends on the exact position of the firm on the labor supply curve, i.e. the exact value of w_i or L_i, the magnitude of this distortion responds endogenously to firm-level demand shocks. Although being useful in guiding measurement and clarifying the basic properties of distortions, the setup in this section is agnostic about what particular sources give rise to the endogenous labor market power distortion of firms and thus, provides little guidance on how international trade might affect such firm-level distortion.

2.2 A Model of Trade and Endogenous Labor Market Distortion

This section develops a partial equilibrium model of heterogeneous firms with oligopsonistic labor markets and international trade. The goal is to build intuition for the source that gives rise to the firm-level endogenous distortion, and for my subsequent empirical analysis of the impact of a trade policy reform. A tradable goods sector is populated by a continuum of Home (H) firms and Foreign (F) firms, indexed by their productivity z, producing differentiated goods.[14] All Home firms are allocated to a continuum of symmetric local labor markets indexed by n. A unique feature of this model is that firms are small within the sector, but are large within a local labor market. When embedding trade into the model, I assume that the Home firms only sell in their domestic market and compete with the Foreign exports in this market. I also focus solely on a one-sided trade policy liberalization of the Home country to obtain sharp predictions on how Home firms respond to trade policy shocks in the labor market.

Utility Function Product market demand and labor market supply are derived from the utility function of a representative household of the economy. This utility function is specified as follow:

$$U = \mathbf{C} - \frac{1}{\bar{\phi}^{\phi}} \frac{\mathbf{L}^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}}, \tag{4}$$

where $\mathbf{C}$ is the sectoral consumption index that increases utility of the household, while $\mathbf{L}$ is the sectoral labor supply index that generates disutility to the household. $\phi > 0$ is the aggregate Frisch elasticity of labor supply. The household maximizes consumption subject to an exogenous budget $\mathbf{I}$ and minimizes disutility from work. The consumption index $\mathbf{C}$ is a CES aggregator of the firm-level demand $c(z)$ within the sector, similar to Melitz (2003),

[14] In this section, since the productivity index z uniquely identifies each individual firm, I replace the subscript i of firm from the previous section by the productivity index z.

Atkeson and Burstein (2008), and Edmond, Midrigan and Xu (2015):

$$\mathbf{C} = \Big[\int_{\Omega^H} c^H(z)^{\frac{\gamma-1}{\gamma}} dz + \int_{\Omega^F} c^F(z)^{\frac{\gamma-1}{\gamma}} dz\Big]^{\frac{\gamma}{\gamma-1}}. \tag{5}$$

In equation (5), $\gamma > 1$ is the constant elasticity of substitution in demand for products across firm z. Ω^H and Ω^F are respectively the mass of active Home firms and Foreign firms in the Home market.

The labor supply index $\mathbf{L}$ is a multi-location nested CES aggregator, a modeling technique I adopt from the recent labor economics literature (Berger, Herkenhoff and Mongey (2019)). More specifically, the representative household allocates labor supply to each location n such that:

$$\mathbf{L} = \Big[\int_{N^H} \mathbf{L}_n^{\frac{\theta+1}{\theta}} dn\Big]^{\frac{\theta}{\theta+1}}. \tag{6}$$

In equation (6), $\theta > 0$ is the constant elasticity of substitution of labor supply across labor markets indexed by n. N^H is the mass of local labor markets within the Home country. Furthermore, within each location n, labor supply is allocated across a finite number of firms K_n so that $\mathbf{L}_n$ can be decomposed as:

$$\mathbf{L}_n = \Big[\sum_{z\in Z_n} L(z)^{\frac{\eta+1}{\eta}}\Big]^{\frac{\eta}{\eta+1}}, \tag{7}$$

where $\eta > 0$ is the labor supply elasticity of substitution across firms z within a local labor market n. Z_n is the productivity set of all active firms in the local labor market n, with the cardinality $|Z_n| = K_n$.[15] I assume that $\eta > \theta$, which implies that firms are closer alternatives within a location, as compared to firms across locations, in the representative household's perspective. Berger, Herkenhoff and Mongey (2019) provide a micro-foundation for the aggregate labor supply system specified in equations (6) and (7), based on a discrete choice model where each individual worker makes labor supply decision to each firm to maximize his(her) utility.[16] Their argument is similar to one employed in the product market to justify the aggregate CES demand system, and is recently used elsewhere in the labor economics literature as in Card et al. (2018). The structures in (4), (5), (6), and (7) are now sufficient

[15]From this setup, the mass of Home firms would be $\Omega^H = \int_{N^H} K_n dn$. For symmetric local labor markets, $K_n \equiv K$, $Z_n \equiv Z$ for all n, and hence $\Omega^H = KN^H$.

[16]The labor supply system in this paper and in Berger, Herkenhoff and Mongey (2019) could be micro-founded from either a static or dynamic discrete choice framework for each individual worker. As shown in Berger, Herkenhoff and Mongey (2019), each worker has random preferences for working at a particular firm, and the elasticity parameters η and θ govern the distribution of this random preferences, conditional on the wage offers by the firms. This micro-foundation approach is used widely for the product demand system that also gives rise to the nested-CES demand in equation (5) (see also Anderson, De Palma and Thisse (1987), Verboven (1996)).

to derive the product demand and labor supply facing each firm.

Product Demand From the aggregate demand system in equation (5), the demand function facing each firm z is:

$$c(z) = p(z)^{-\gamma}\mathbf{P}^{\gamma-1}\mathbf{I}, \tag{8}$$

where $\mathbf{P}$ and $\mathbf{I}$ are respectively the exogenous aggregate price index and aggregate income spent on the sector. Firm z takes the aggregate price index $\mathbf{P}$ as given in its optimization problem because it is small within the sector, whereas the aggregate expenditure $\mathbf{I}$ depends on the broader structure of the economy and is assumed to be predetermined. The aggregate price index can be shown to have the following form:

$$\mathbf{P} = \Big[\int_{\Omega^H} p^H(z)^{1-\gamma}dz + \int_{\Omega^F} p^F(z)^{1-\gamma}dz\Big]^{\frac{1}{1-\gamma}}. \tag{9}$$

Production Technology The Home firms only produce and sell in the domestic market. Firm has productivity level z, incurs a fixed cost f in terms of a numeraire good, and uses labor as the only factor of production to produce output $y(z)$:

$$y(z) = c^H(z) = zL(z), \tag{10}$$

where $L(z)$ is the labor factor use in production of firm z. The presence of the fixed cost f allows for an endogenous form of entry and exit, which will be the main mechanism through which trade policy reform affects the labor market equilibrium in this model.

Labor Supply From the aggregate labor supply system in equations (6) and (7), the labor supply function facing each firm z, located in labor market n, can be derived as:

$$L(z) = \bar{\phi}w(z)^{\eta}\mathbf{W}_n^{\theta-\eta}\mathbf{W}^{\phi-\theta}, \tag{11}$$

where $\mathbf{W}_n$ $\forall n$ is a local labor market wage index, specified as:

$$\mathbf{W}_n = \Big[\sum_{z\in Z_n} w(z)^{1+\eta}\Big]^{\frac{1}{1+\eta}} \tag{12A}$$

and $\mathbf{W}$ is an aggregate labor supply shifter of the sector at the national level in the Home country:

$$\mathbf{W} = \Big[\int_{N_H} \mathbf{W}_n^{1+\theta}dn\Big]^{\frac{1}{1+\theta}}. \tag{12B}$$

Since local labor markets have measure zeros in the national economy, $\mathbf{W}$ is exogenously given to each firm. However, in contrast to the product market, because firms are large within a local labor market, the local labor market wage index $\mathbf{W}_n$ is endogenous from firm z's perspective. Due to this particular feature of the model, firms exhibit a strategic distortion in the local labor market. In other words, firm z's wage offer $w(z)$ (or employment level $L(z)$) affects the aggregate local labor market wage (employment) index.

Firm-level Equilibrium and Endogenous Distortion A Home firm z chooses its price $p(z)$ and wage $w(z)$ to solve for the following profit maximization problem:

$$\Pi(z) = p(z)c(z) - w(z)L(z) - f, \tag{13}$$

where each endogenous variable $p(z), c(z), w(z), L(z)$ is subject to the constraints given by equations (8)-(11).[17] Given these contraints, the firm's problem in (13) is to decide either the optimal employment $L(z)$ or wage $w(z)$ level. Taking the FOC, I obtain the following expression for the endogenous labor market distortion of firm z:

$$\tilde{\chi}^e(z) = 1 + \chi^e(z) = \frac{MRPL(z)}{w(z)} = (1 - s(z))\Big(1 + \frac{1}{\eta}\Big) + s(z)\Big(1 + \frac{1}{\theta}\Big). \tag{14}$$

In equation (14), s_z is the wage-bill share of firm z within the local labor market:

$$s(z) = \frac{w(z)L(z)}{\sum_{z' \in Z_n} w(z')L(z')} = \frac{w(z)^{\eta+1}}{\sum_{z \in Z_n} w(z')^{\eta+1}}. \tag{15}$$

Equations (14)-(15) provide key intuition for the sources of the endogenous distortion $\tilde{\chi}^e(z)$. In particular, $\tilde{\chi}^e(z)$ depends on two key parameters: within-market (η) and across-market (θ) elasticity of substitution of labor supply, and firm's own local labor market share $s(z)$. When the firm accounts for an infinitesimal share of the local labor market such that $s(z) \to 0$, the endogenous distortion reach the lower-bound of $(1 + \frac{1}{\eta})$. On the other hand, a monopsonist employer with $s(z) = 1$ incurs a distortion with the magnitude of $(1 + \frac{1}{\theta})$, the upper-bound of the distortion in this model.[18]

[17] To focus on the endogenous distortion, notice that I have set $\chi^x = 0$ in the firm's problem in equation (13), as compared to the problem in equation (1).

[18] Recall that η is assumed to be greater to θ. Therefore, $\frac{1}{\theta} > \frac{1}{\eta}$. When all firms are infinitesimally small within the local labor market, the distortion converges to the constant $\frac{1}{\eta}$, which is equivalent to the case of monopsonistic competition in the labor market (see also Ranjan and Rodriguez-Lopez (2019)).

Entry Game and Market Equilibirum I allow for the endogenous entry and exit of firms within the local labor market in response to aggregate sectoral trade shocks. This is motivated by the observed empirical patterns that trade shocks induce endogenous entries and exits, and in turns, affect employer concentration within local labor markets. The entry game follows closely the modeling approach for oligopoly models in international trade, such as those in Atkeson and Burstein (2008), Eaton, Kortum and Sotelo (2012), Edmond, Midrigan and Xu (2015), and most recently Gaubert and Itskhoki (2021). The only difference in this model is that firms compete strategically in the labor market rather than in the product market.

To start with, I assume that the Home firms within a local labor market play a static Bertrand game of wage competition, i.e., firms choose the optimal wage level so as to maximize their profit, internalizing the effect of its own action and the action of other local firms on the local labor market.[19] Each perfectly symmetric local labor market n has access to an identical integer number of potential firms, with productivity being ranked as:

$$z^{(1)} > z^{(2)} > z^{(3)} > ... > z^{(k)} > ... \tag{16}$$

I focus on the equilibrium where firms make sequential entry decisions based on the decreasing order of their productivity ranking. In particular, the most productive firm $z^{(1)}$ makes the entry decision first, followed by the second-most productive firm $z^{(2)}$, and so on. When making entry decisions, each firm can compute perfectly what its profit would be, knowing that all the more productive players have already entered the market. Firm $z^{(k)}$ decides to operate in the market as long as its profit is greater or equal to zero:

$$\Pi^K(z^{(k)}) \geq 0. \tag{17}$$

Notice that in equation (17), the profit function now has a superscript K. The sole purpose of this superscript is to explicitly indicate that the number of active firms in the local labor market enters the profit calculation of the firm $z^{(k)}$. Proposition 1 below defines the equilibrium of the model, its existence as well as its uniqueness.

Proposition 1 *An equilibrium in the environment set up by equations* (5)-(11)*, in which firms make sequential entry decisions based on the decreasing order of their productivity ranking, is fully determined by the equilibrium number of active firms K^* in each local labor market. A unique equilibrium K^* exists such that no firm has incentives to enter or exit the*

[19] I use Bertrand competition in this model as it is more natural to my subsequent proofs. However, in this environment, it can be shown that Cournot competition provides an identical outcome for the model, as being used in Berger, Herkenhoff and Mongey (2019) without entry.

market. In the equilibrium with K^* firms, the least productive firm operating in the market is $z^{(K^*)}$.

Proof. *See the Theory Appendix.*

Proposition 1 states that the number of firms K^* sufficiently characterizes a market equilibrium, and that a unique equilibrium with K^* firms within each local labor market exists. This unique equilibrium arises from the fact that firms are required to make sequential moves by their productivity ranking, and from a so-called *profit monotonicity* condition, specified as follows:

$$\Pi^K(z^{(k)}) \geq \Pi^{K+1}(z^{(k)}) \geq \Pi^{K+1}(z^{(k+1)}). \tag{18}$$

Intuitively, equation (18) states that more productive firms always have higher profit than less productive firms given any market conditions (the latter inequality). Not only so, more productive firms earn higher profit if there are fewer active firms, i.e., less competition, in the local labor market (the former inequality). As a result, if one were to observe that the firm $z^{(k+1)}$ operates in the market, it must be true that the firm $z^{(k)}$ also operates in that market. The *profit monotonicity* condition (18) allows me to solve for the equilibrium using backward induction, and to show that a unique equilibrium K^* exists.[20]

Market Equilibirum with Trade Policy Trade policy is modeled in this environment using two instruments: the output tariff (τ^O), and the input tariff (τ^I). The output tariff is the import tariff imposed directly on the Foreign product $c^F(z)$ sold in the Home market. The input tariff, on the other hand, is the tariff imposed on imported intermediate inputs used by Home firms in the production process. I first focus on the impact of the output tariff on the market equilibrium, and then explore the implication of the input tariff when the production function involves an intermediate input, which requires a slight modification of the production function in equation (10).

Trade shocks working through changes in the output tariff (τ^O) transmit their competitive pressure to the aggregate price index, i.e. $\mathbf{P}(\tau^O)$. This price index in turn shifts the labor demand of firms, and consequently affects the local labor market equilibrium. To see this, in equation (9), I assume that the mass of Foreign firms selling in the Home market Ω^F are

[20] As in oligopoly-type models, there are multiple equilibria in this environment if the entry game is specifed differently. However, these equilibria are often intractable and uninteresting. By requiring firms to make sequential moves in a particular order, I can turn attention to an equilibrium that is most informative. Edmond, Midrigan and Xu (2015) shows in their quantitative exercise that the exact ordering of moves matters little in practice. The term *profit monotonicity* is coined by Eaton, Kortum and Sotelo (2012) when describing the equilibrium in their oligopoly model.

subject to an ad-volerum tariff τ^O such that the price received by the Foreign firms, denoted by $p^{F*}(z)$, is a fraction of the Home market price $p^F(z)$:

$$p^F(z) = (1 + \tau^O)p^{F*}(z). \tag{19}$$

To simplify the model, I also assume that the Home market is small enough so that Ω^H is an order of magnitude smaller than Ω^F. This assumption has two implications. First, Ω^F can be held fixed, and $p^{F*}(z)$ does not respond to changes in the Home market environment. Second, when there are entries and exits of firms in the Home market in response to tariff changes, this assumption guarantees a monotonic movement in the aggregate price as a function of tariff. It is straight-forward to rewrite equation (9) in the following form and show that the aggregate price index $\mathbf{P}$ is an increasing function of the output tariffs τ^O:

$$\mathbf{P}(\tau^O) = \Big[\int_{\Omega^H} p^H(z)^{1-\gamma} dz + (1 + \tau^O)^{1-\gamma} \int_{\Omega^F} p^{F*}(z)^{1-\gamma} dz \Big]^{\frac{1}{1-\gamma}}, \tag{20}$$

where $\mathbf{P}'(\tau^O) \geq 0$. The labor demand curve in this model, i.e. the $MRPL(z)$ curve, can be derived as:

$$MRPL(z) = z^{\frac{\gamma-1}{\gamma}} L(z)^{-\frac{1}{\gamma}} \mathbf{P}(\tau^O)^{\frac{\gamma-1}{\gamma}} \mathbf{\Xi}, \tag{21}$$

where $\mathbf{\Xi} = (\frac{\gamma-1}{\gamma})\mathbf{I}^{\frac{1}{\gamma}} > 0$ is an aggregate constant. As can be seen from equations (20)-(21), tariff changes affect the aggregate price index $\mathbf{P}(\tau^O)$ and shift the $MRPL(z)$ curve. Firms observe these changes in the aggregate environment, re-calculate their profit taking into account the competition structure in the labor market, decide if they should operate, and if operating, the optimal level of wage $w^*(z)$.

From proposition 1 and starting from the equilibrium with a high-level output tariff, there exists a "cut-off" firm which represents the lowest productivity firm operating in the market, i.e. $z^{(K^*)}$. When the output tariff is lowered, competitive pressure drives down each Home firm's profit. This makes the operating decision of low productivity firms near the "cut-off" less profitable, for example ... $z^{(K^*-1)}$, $z^{(K^*)}$, and induces exit of these firms. As a result, the equilibrium number of firms decreases, local labor market shares are reallocated towards surviving firms, and the average distortion increases in the local labor market.[21] I summarize the implications of a change in the output tariff on the local labor market equilibrium with proposition 2.

[21] This mechanism is similar to the canonical Melitz (2003) model. However, in Melitz (2003), there is no labor market distortion. Ranjan and Rodriguez-Lopez (2019) allow for labor market distortion in Melitz (2003)'s environment, but assume such distortion to be constant, and thus, not responding to competitive trade shocks.

Proposition 2 ***(Equilibrium with Output Tariff)*** *Under the environment set up by equations (5)-(11), in which firms make sequential entry decisions based on the decreasing order of their productivity ranking,* ***lowering*** *ouput tariffs (τ^O) induces exit of less productive firms, reallocates local labor market shares towards more productive firms, and* ***increases*** *the average endogenous distortion in the local labor market. Formally, $K^{*'}(\tau^O) \geq 0$, and for $\forall z \geq z^{(K^*)}$, $s'(\tau^O, .) \leq 0$, $\tilde{\chi}^{e'}(\tau^O, .) \leq 0$.*

Proof. *See the Theory Appendix.*

To consider the effect of an input tariff change, I need to modify the production function in equation (10) to involve an intermediate input. Specifically, the modified production function is as follow:

$$y(z) = c^H(z) = zL(z)^{\alpha} M(z)^{1-\alpha}, \tag{22}$$

where $M(z)$ is the intermediate input used by firm z, and α is the factor share of labor. Firms still need to incur a fixed cost f in terms of a numeraire good to produce goods. The price of intermediate input is determined by the world market, i.e. perfectly competitive, and subject to the Home country's input tariff (τ^I). More formally:

$$p_M = p_M^{World}(1 + \tau^I). \tag{23}$$

From the equations (22)-(23), lowering the input tariff has an intuitive effect on the firm-level labor demand. Specifically, lowering the input tariff induces firms to use more intermediate input, which through the production function, increases the marginal revenue product of each worker. Therefore, in contrast to the impact of lowering the output tariff, lowering the input tariff decreases production costs, drives up profit, and induces entries of firms that ex-ante has productivity below the "cut-off" firm, for example $z^{(K^*+1)}$, $z^{(K^*+2)}$,... Furthermore, the magnitude of the effect by the input tariff is magnified by a factor of $\frac{1-\alpha}{\alpha}$, which is the ratio of factor shares between labor and intermediate input. The impact of the input tariff on market equilibrium is summarized by proposition 3.

Proposition 3 ***(Equilibrium with Input Tariff)*** *Under the environment set up by equations (5)-(11) and with modifications in equations (22)-(23), in which firms make sequential entry decisions based on the decreasing order of their productivity ranking,* ***lowering*** *input tariffs (τ^I) induces entry of less productive firms, reallocates local labor market shares towards new entrants, and* ***decreases*** *the average endogenous distortion in the local labor market. Formally, $K^{*'}(\tau^I) \leq 0$, and for $\forall z \geq z^{(K^*)}$, $s'(\tau^I, .) \geq 0$, $\tilde{\chi}^{e'}(\tau^I, .) \geq 0$. Furthermore, compared to the output tariff, the impact of the input tariff is magnified by $\frac{(1-\alpha)(\frac{\gamma-1}{\gamma})}{1-(1-\alpha)(\frac{\gamma-1}{\gamma})}$,* *the adjusted relative factor share between labor and intermediate input.*[22]

Proof. *See the Theory Appendix.*

Proposition 3 concludes my theoretical analysis. Before moving on to the empirical analysis, it is important to point out that all the predictions in propositions 1-3 still hold true if the product market is assumed to be perfectly competitive.

3 Measuring Labor Market Distortion

This section develops a production-based framework to estimate the *overall distortion* at the firm-level. As will be clear shortly, identification of the production function, and hence labor market distortion, comes from the plausible assumptions of firm's decision making, rather than any structural assumptions on either the product or the labor market. Section 3.1 sketches out an estimation approach based on a nonparametric method, building on the recent work by Gandhi, Navarro and Rivers (2020) on production function estimation. Section 3.2 briefly describes the data on Chinese manufacturing firms used in estimation, and section 3.3 presents the main empirical results of measured distortion.

3.1 Estimating Distortion from Production Data

My estimation strategy of the labor market distortion follows directly from equation (3). To simplify notation, I omit the subscript i when it does not cause any confusion. It is convenient to rewrite the expression of the distortion in equation (3) as a ratio between the revenue elasticity of labor and the wage-bill share of total revenue as follows:

$$\tilde{\chi} \equiv \frac{MRPL}{w} = \frac{\frac{\partial R(.)}{\partial L}}{w} = \frac{\frac{\partial r(.)}{\partial l}\frac{R}{L}}{w} = \frac{\frac{\partial r(.)}{\partial l}}{\frac{wL}{R}}, \tag{24}$$

where r and l, respectively, are the natural logs of total revenue and labor factor.[23] Denoting the revenue elasticity of labor as $\theta^L \equiv \frac{\partial r(.)}{\partial l}$, and the wage-bill share of total revenue as $\alpha^L \equiv \frac{wL}{R}$, the distortion $\tilde{\chi}$ in equation (24) could now be expressed in the following short-form:

$$\tilde{\chi} = \frac{\theta^L}{\alpha^L}. \tag{25}$$

[22]This adjusted relative factor share converges to $\frac{1-\alpha}{\alpha}$ as $\gamma \to \infty$ i.e. perfect competition in product market.

[23]Notice that the wage-bill share of total revenue here (α^L) is the expenditure share on labor within each firm, and distinct from the local labor market wage-bill share defined in section 2, which aims to measure the labor market share of each firm within a location.

Since information about the wage-bill is readily available in most production datasets, the task now is to estimate the revenue elasticity with respect to labor (θ^L) from a revenue production function. To achieve this goal, I begin by specifying a revenue production function of firm in the log-form as follows:

$$r_t = f(k_t, l_t, m_t) + \omega_t + \varepsilon_t, \tag{26}$$

where r_t, k_t, l_t, m_t are the natural logs of revenue, capital, labor, and material. ω_t measures the revenue productivity, i.e. revenue TFP, in period t, and ε_t is a random measurement error. Here, $f(.)$ is a revenue production function, and allowed to be nonparametric.[24]

Identification and Estimation of Production Function To identify and estimate the revenue production function specified in equation (26), I build on a recent nonparametric estimation method proposed by Gandhi, Navarro and Rivers (2020), henceforth, the GNR method. As is common in the productivity literature, identification of the production function in the GNR method is rooted in the timing assumptions in decision making by the firm. However, in constrast to other existing methods such as those proposed by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves and Frazer (2015), GNR exploits an additional restriction derived from profit-maximizing behavior of the firm with respect to materials to identify the revenue elasticity of this input. In what follows, I provide a brief description of how I adapt the GNR method to identify and consistently estimate the revenue elasticity of labor, θ^L. A distinct feature of my approach as compared to the original GNR method is that I do not need to assume perfect competition in the product market, since my goal is to identify a revenue production function.

My estimation procedure is implemented in two stages. In the first stage, firm's profit-maximizing behavior with respect to material is exploited to provide identification information for the revenue elasticity of material, i.e. $\frac{\partial r(.)}{\partial m}$. The intuition is that when firms maximize profit with respect to factor inputs, revenue elasticities have to be equal to expenditure shares for all factors that are *not* subject to market frictions. In this case, I assume that the market for material is relatively frictionless, and hence, material expenditure share is informative about the revenue elasticity of this factor.[25] Following GNR, in the first-stage,

[24]The specification of the (log) revenue production function in equation (26) could be microfounded within a large class of demand structures that dictate the firm-specific price as a power function of quantity. See De Loecker (2011) for an example.

[25]In principle, material could be subject to market frictions as well. To alleviate the concerns about frictions in this market, I control for an extensive set of exogenous state variables that could affect the material demand decisions. Therefore, as long as firms do not possess market power in the market for material, estimating its revenue elasticity from expenditure share would be consistent. This approach is also used in other empirical work, for example in Halpern, Koren and Szeidl (2015).

I estimate the following share-regression using a nonlinear least-square (NLS) procedure:[26]

$$log(s_t^M) = log\frac{\partial}{\partial m_t}f(k_t, l_t, m_t) - \varepsilon_t. \tag{27}$$

In equation (27), s_t^M is the expenditure share of material obtained directly from the data, and is defined as $s_t^M = \frac{p_t^M M_t}{R_t}$. The nonparametric elasticity function $\frac{\partial f(.)}{\partial m_t}$ is approximated by a second-order polynomial sieve. The estimation of equation (27) provides me with two ouputs to use in the second stage: the revenue elasticity of material $\frac{\partial \hat{f}(.)}{\partial m_t}$, and the random shock $\hat{\varepsilon}_t$.

In the second stage, the production function is fully indentified using a Generalized Method of Moments (GMM) procedure. Specifically, given the estimate of $\frac{\partial f(.)}{\partial m_t}$ and by simple integration, production function $f(.)$ is identified up to a constant $C(.)$ as a function of k_t, l_t. This integration is denoted by $D^\varepsilon(k_t, l_t, m_t)$:

$$\int \frac{\partial}{\partial m_t}f(k_t, l_t, m_t)dm_t = f(k_t, l_t, m_t) + C(k_t, l_t) \equiv D^\varepsilon(k_t, l_t, m_t). \tag{28}$$

Plug the expression in equation (28) back to the original specification of production function in equation (26), I can rewrite the productivity term as:

$$\omega_t = (r_t - \varepsilon_t - D^\varepsilon(.)) + C(k_t, l_t). \tag{29}$$

Following the productivity literature, firm productivity is assumed to follow a flexible Markov process:

$$\omega_t = h(\omega_{t-1}) + \eta_t, \tag{30}$$

where η_t is an exogenous productivity shock to the firm at time t. Importantly, the exogeneity assumption imposed here is that k_t and l_t are predetermined, and do not respond to η_t. In other words, I assume that capital and labor factors are subject to planning, and chosen based solely on the information about the expected productivity captured by $h(\omega_{t-1})$. The only factor that responds to the productivity shock η_t is the material m_t, the elasticity with respect to which is already identified in the first stage. The Markov productivity process in equation (30) provides exclusion restrictions needed to identify the function $C(.)$. Let's denote $\Psi_t \equiv r_t - \varepsilon_t - D^\varepsilon(.)$, and combine equations (29)-(30), I can now rewrite the Markov productivity process as:

$$\Psi_t = -C(k_t, l_t) + h(\Psi_{t-1} + C(k_{t-1}, l_{t-1})) + \eta_t. \tag{31}$$

[26]The derivation of this share-regression from firm's FOC is provided in the Appendix.

Equation (31) nonparametrically identifies $C(.)$ and $h(.)$, and in turn, provides identification of the revenue production function. Estimation of equation (31) is performed using a GMM procedure.[27] In my estimation, other than primary factors such as capital and labor, I also control for a vector of state variables that may affect input demand decision of firms, including year, location and industry fixed effects, firm's ownership type, export status, and tariff levels associated with the firm's industry.

Compute the Distortion Given estimates from the revenue production function, I can now compute the empirical measure of the labor market distortion expressed in equation (25). Since ε_t is a random measurement error, and does not affect firm's labor demand decision, I need to correct for this term in calculating the *expected* revenue that enters the denominator of the distortion in equations (24)-(25). The estimation of equation (27) in the first stage does provide me with an estimate of the measurement error, i.e. $\hat{\varepsilon}_t$. The measure of the distortion, therefore, can be computed as:

$$\tilde{\chi} = \frac{\hat{\theta}^L}{\hat{\alpha}^L} = \frac{\frac{\partial \hat{r}(.)}{\partial l}}{\frac{wL}{R} \times \exp(\hat{\varepsilon}_t)}. \tag{32}$$

This final step concludes my estimation procedure for the *overall* labor market distortion, which is based solely on the production function estimation approach.

3.2 Chinese Firm-level Data

The Chinese firm-level data comes from China's Annual Survey of Industrial Enterprises (ASIE) from 1998 to 2007. This dataset is a rather standard panel dataset covering all industrial private firms with sales above 5 million Renminbi (RMB) and all state-owned enterprises (SOEs) . The dataset encompasses more than 90% of industrial activities in China in terms of gross output during the sample period (Brandt, Biesebroeck and Zhang (2014)). Table D1 in the Appendix reports the main aggregate statistics of this dataset, which matches with published official statistics from the China's National Bureau of Statistics, and confirms the dataset's quality. The dataset contains all variables required for the production function estimation, including total gross output (revenue), capital stock, employment, material (in monetary values). In addition, the dataset also contains information about wage-bill, ownership status, export, detailed geographical code, and firms' four-digit industry affiliation. My final cleaned sample consists of $1,235,801$ firm-year observations, spanning over 10 years, and 422 four-digit industries. Throughout this paper, the unit for

[27] See details of this GMM procedure in section 6 of the Appendix.

the local labor market in Chinese firm-level dataset is prefecture, which is the lowest-ranked administrative unit that has authority to set labor market regulations such as the minimum wage policy and the *hukou* household registration status. This administrative unit of the local labor market has also been discussed and used in Hau, Huang and Wang (2020), Tombe and Zhu (2019). The final firm-level data cover 461 Chinese prefectures throughout the sample period. Working with Chinese firm-level data requires careful data filtering process. I elaborate on this process in the section C of the Appendix.

3.3 Empirical Measure of Distortion

Table 1 reports the empirical results for the revenue elasticites and labor market distortion across 29 two-digit Chinese manufacturing industries. Since my production function is nonparametric, I can recover the distribution of each revenue elasticity and the firm-level distortion within each industry. Across all industries, my estimation procedure's performance is remarkably stable and produces an average capital elasticiy of 0.07, an average labor elasticity of 0.09, and an average material elasticity of 0.74. The average revenue return to scale (RTS), is 0.90.[28] The average magnitude of the labor market distortion $\tilde{\chi}$ estimated for China's whole manufacturing sector is 2.14, implying an average *overall* pass-through rate of 47% of an idiosyncratic demand shock to wage. This pass-through rate suggests that, for instance, of a productivity shock that increases marginal revenue product of a worker by one dollar, only 47 cents is shared with the worker in the form of wage payment. The median value of estimated distortion is 1.49, indicating that the distribution of firm-level distortion is skewed to the right. Across all industries, both the mean and median of the distortion are consistently greater than one. This empirical fact suggests that Chinese manufacturing firms face pervasive frictions in the labor market during the 1998-2007 period.

Labor Market Distortion across Industries and Years Given these estimates of the firm-level labor market distortion, I can now investigate its distribution as well as its correlation patterns across industries and years. The left panel of Figure 2 illustrates the distribution of $log(\tilde{\chi}_i)$ for the whole sample. As one can see, the log of this distribution well on the right of zero (no distortion) threshold. The right panel of Figure 2 displays the evolution of labor market distortion distribution over three equidistant years within my

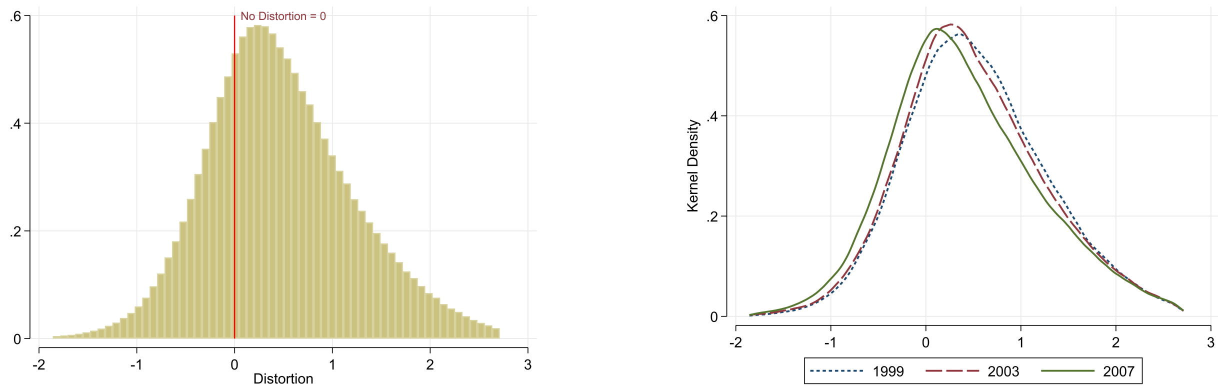
[28]If one is willing to impose constant RTS of *physical* production to the whole Chinese manufacturing sector, this revenue RTS implies the average markup of 1.11 (or 11%) for the sector. This approach has been developed by Flynn, Gandhi and Traina (2019) to estimate the markup for the US, with the magnitude of US markup in the similar range to what I obtain for China. In section 5, I adopt this approach to measure Chinese firm-level markups.

Table 1: Revenue Elasticities and Labor Market Distortion by Industry

Industry	Capital	Labor	Material	RTS	$\tilde{\chi}_i$ Mean	$\tilde{\chi}_i$ Median	No. Obs
13. Food Processing	0.06	0.08	0.74	0.88	3.82	3.09	82003
14. Food Production	0.07	0.08	0.72	0.88	1.99	1.39	27268
15. Beverage	0.09	0.09	0.69	0.87	2.64	2.01	19511
16. Tobacco	0.15	0.09	0.68	0.92	2.51	1.60	300
17. Textile	0.06	0.09	0.76	0.91	1.92	1.51	121960
18. Garments	0.05	0.13	0.74	0.92	1.69	1.26	60683
19. Leather	0.05	0.12	0.75	0.91	1.94	1.32	30338
20. Timber	0.07	0.10	0.73	0.90	2.35	1.83	26268
21. Furniture	0.05	0.12	0.74	0.91	2.26	1.65	15096
22. Paper-making	0.07	0.09	0.75	0.91	2.34	1.89	43534
23. Printing	0.10	0.08	0.72	0.90	1.49	1.12	19427
24. Cultural	0.06	0.11	0.74	0.92	1.54	1.15	16558
25. Petroleum Processing	0.09	0.09	0.73	0.91	3.89	2.95	9108
26. Raw Chemical	0.08	0.07	0.74	0.88	2.39	1.72	96338
27. Medical	0.11	0.11	0.67	0.89	2.87	2.09	27221
28. Chemical Fibre	0.07	0.08	0.78	0.94	2.70	2.12	4848
29. Rubber	0.07	0.08	0.73	0.88	1.72	1.19	14630
30. Plastic	0.07	0.08	0.75	0.91	1.97	1.48	57862
31. Nonmetal Products	0.08	0.08	0.71	0.87	1.40	1.03	119803
32. Processing of Ferrous	0.08	0.11	0.75	0.94	4.39	3.62	27343
33. Processing of Nonferrous	0.07	0.09	0.76	0.91	3.62	2.80	17726
34. Metal Products	0.07	0.08	0.75	0.89	1.66	1.22	61971
35. Ordinary Machinery	0.08	0.08	0.73	0.89	1.71	1.25	92733
36. Special Equipment	0.07	0.09	0.71	0.88	1.91	1.31	44527
37. Transport Equipment	0.08	0.10	0.73	0.91	1.97	1.44	55016
39. Electric Machinery	0.07	0.08	0.75	0.91	2.09	1.45	73088
40. Electronic and Telecom	0.08	0.11	0.72	0.92	1.96	1.41	34803
41. Measuring Instruments	0.07	0.10	0.71	0.88	1.61	1.20	15190
42. Art Work	0.05	0.10	0.73	0.88	1.54	1.10	20648
All Industry	0.07	0.09	0.74	0.90	2.14	1.49	1,235,801

Note: The table reports estimates of the revenue elasticities of factors (capital, labor, material), the revenue return to scale (RTS), and the measured *overall distortion* ($\tilde{\chi}_i$) from production function estimation in section 3. Except for the distortion, all other statistics are the mean of respective distributions. The table trims observations above and below the 0.5^{th} and 99.5^{th} percentiles. The last column reports the number of observations for each two-digit industry. Notice that the RTS would not be equal to 1 in this case because it contains markups. If I impose the constant RTS assumption of the physical production function, the average markup of Chinese manufacturing sector is 11%.

Figure 2: Distribution of the Labor Market Distortion ($log(\tilde{\chi_i})$)



Note: The figure illustrates the histogram (left) and kernel density (right) of the log measured labor market distortion ($\tilde{\chi_i}$) from production function estimation. The left panel shows the distribution of distortion across all firm-year observations. No distortion cutoff is where $log(\tilde{\chi_i}) = 0$. The right panel displays the evolution of distortion distribution over three equidistant years: 1999, 2003, 2007.

sample period: 1999, 2003 and 2007. As shown in the figure, across the three years, the distribution of distortion has shifted to the right, with decreases in both the mean and median. Furthermore, the dispersion of distortion distribution has also reduced substantially.[29]

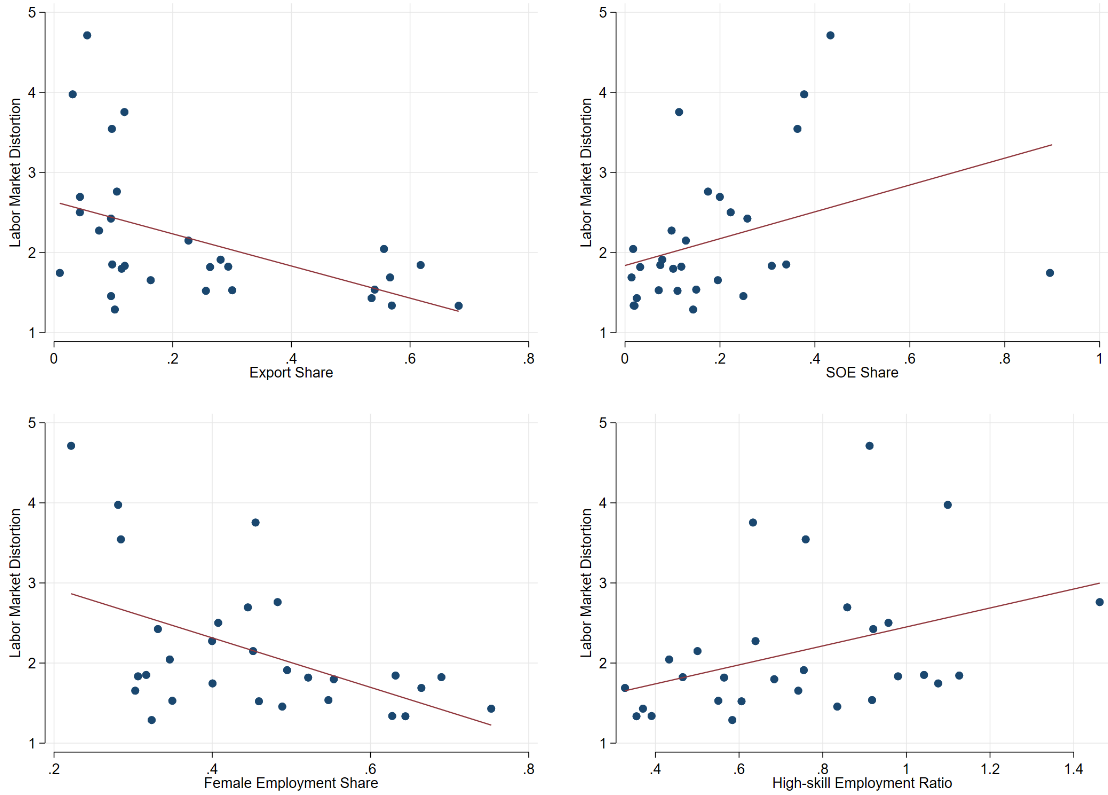
In Figure 3, I show the patterns of correlation between the distortion and several two-digit industry characteristics which might potentially be associated with the degree of labor market distortion, including industry's export share, state ownership share, high-skill employment ratio, and female employment share.[30] The data for industry-level characteristics is extracted from the firm-level data for the year 2004, when more detailed information is reported for each firm.[31] The top panels of Figure 3 illustrate that more export-oriented industries tend to have lower levels of labor market distortion, while industries with higher shares of state ownership exhibit higher levels of distortion. In the bottom panels, industries that employ more female workers tend to have lower distortion, while the reverse is true for industries with larger high-skill employment ratio.

[29]The reduction in the dispersion of the distortion typicall implies that there might be a lesser degree of labor market misallocation over time. This rationale is applied elsewhere in the misallocation literature, for example as in Hsieh and Klenow (2009), Lu and Yu (2015), Morlacco (2020).

[30]The high-skill employment ratio is defined as the ratio between the number of workers that finish high-school and the number of workers that finish only secondary-school. The correlation patterns are also robust to using finishing college degree as an alternative measure of skill level.

[31]The year 2004 is a China's Census year, therefore, more detailed data is collected for each firm (see also Brandt, Biesebroeck and Zhang (2014)).

Figure 3: Labor Market Distortion ($\tilde{\chi}_i$) and Industry Characteristics (in 2004)



Note: The figure illustrates the correlations between the measured labor market distortion ($\tilde{\chi}_i$) and (2-digit) industry characteristics in 2004. The industry characteristics include: export share, state ownership (SOE) share, female employment share and high-skill employment ratio. High-skill employment ratio is defined as the ratio between high-school and secondary-school degree workers. The figure is based on the data in 2004, because this is the only year that the employment composition information is available.

Labor Market Distortion across Firms To examine the correlation patterns of distortion across firms, I correlate firm-level labor market distortions with measured productivity, employment size, export status, ownership status, local labor market concentration, and local minimum wage. As shown in Table 2, more productive firms have higher level of labor market distortion, regardless of the covariates included. Conditioning on productivity, larger firms are associated with less distortion. Columns (3)-(5) show that exporting firms and foreign-invested firms incur less distortion, while state-owned enterprises (SOEs) are more distorted in the labor market.[32]

[32]This is perhaps surprising, given the fact that workers at SOEs typically have lower marginal revenue products (Hsieh and Song (2015)). However, this is consistent with state firms being subject to more frictions in hiring and firing decisions, and in line with empirical results I obtain in section 4, where the labor supply elasticity estimated from an exogenous demand shifter for SOEs is much lower.

Table 2: Labor Market Distortion and Firm Characteristics

Dependent Variable	$\tilde{\chi}_i$	$\tilde{\chi}_i$	$\tilde{\chi}_i$	$\tilde{\chi}_i$	$\tilde{\chi}_i$
	(1)	(2)	(3)	(4)	(5)
Productivity (TFPR)	6.539	6.659	6.715	6.754	7.131
	(0.014)	(0.014)	(0.014)	(0.014)	(0.016)
Employment		-0.317	-0.286	-0.290	-0.353
		(0.002)	(0.002)	(0.002)	(0.002)
Exporting			-0.332	-0.294	-0.192
			(0.003)	(0.003)	(0.004)
Foreign-Owned				-0.222	-0.124
				(0.005)	(0.006)
SOE				0.065	0.081
				(0.006)	(0.007)
HHI (Employer Concentration)					0.120
					(0.006)
Minimum Wage (Monthly)					-0.003
					(0.000)
Observations	1,235,801	1,235,801	1,235,801	1,235,801	985,076
R-squared	0.374	0.396	0.401	0.402	0.431
Fixed Effects					
Industry	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes

Note: The table reports the regression results of the measured distortion ($\tilde{\chi}_i$) on firm-level characteristics. The HHI in column (5) is the Herfindahl-Hirschman Index of employer concentration (measured in wage bill) within a prefecture. Minimum wage data (at prefecture-level) is only available from 2000-2007, thus there are less observations in column (5). Robust standard errors are enclosed in parentheses.

Importantly, local labor market characteristics seem to play an important role in determining the firm-level labor market distortion. Column (5) shows that more concentrated and lower minimum wage labor markets are associated with higher firm-level distortion. This pattern is consistent with the existence of an endogenous source of labor market distortions that arise from local labor market conditions, which I theoretically motivate in section 2, and explore further empirically in section 4.[33]

[33]The HHI in column (5) is the Herfindahl-Hirschman Index of employer concentration (in wage-bill share) within a prefecture, which is the geographic unit of local labor market in this paper.

4 Labor Market Power as Endogenous Distortion

As described in section 2, labor market distortion can be exogenous or endogenous to firm-level shocks. In the sole presence of the exogenous distortion, the firm-level wage does not respond to an idiosyncratic labor demand shock while employment possibly does. This section provides reduced-form evidence that the firm-level wage does respond proportionately along with employment to an exogenous demand shifter, namely the granting of permanent normal trade relations to China by the United States, which was a major trade policy uncertainty shock (henceforth, US-China TPU shock).

By comparing the responses of wage and employment to an exogenous demand shock, I can quantify the magnitude of the endogenous distortion in the labor market. I will show firstly that a firm's responses in terms of wage and employment imply an upward sloping labor supply curve, with an average elasticity of 1.57. This indicates that the magnitude of the endogenous distortion on average accounts for almost 76% of the overall distortion measured by the production function approach in section 3. And secondly, the pattern of responses is dependent on the firm's local labor market wage-bill share, which suggests the existence of a strategic component of firms' noncompetitive conduct in the labor market that is consistent with the theory proposed in section 2.

To credibly identify firms' wage and employment response to an idiosyncratic demand shock, I use the removal of US trade policy uncertainty towards China, which is associated with China's accession to WTO in 2001, as a quasi-experimental demand shifter. The intuition is that the removal of TPU increases the expected profit of entering the export market or expanding export activities, and therefore raises firm-level labor demand. The measurement of TPU and the behavioral response of firms towards TPU removal is modeled and studied by Handley and Limao (2015), Handley and Limão (2017). In the context of China's accession to WTO, Pierce and Schott (2016) estimate the impact of the US-China TPU shock on US manufacturing employment, and finds that the entry of new Chinese exporters increases significantly following the TPU shock. Handley and Limão (2017) investigate the impact of the US-China TPU shock on US's imports from China, US prices and welfare. Most importantly, these studies provide robust evidence that the across-industry variation of the US-China TPU shock is largely exogenous from the perspective of Chinese firms. I follow this literature in measuring the TPU shock, and treat such shock as random. However, different from this literature's focus, my focus is on the *relative* responses of wage and employment to infer endogenous distortions in the labor market.[34]

[34] In Handley and Limao (2015) and Handley and Limão (2017), the labor market is assumed to be perfectly competitive, and hence, there is no room for an empirical investigation of labor market distortions.

Formally, the US-China TPU shock at four-digit industry level is measured by the gap between the "Column 2" tariffs and the MFN tariffs faced by the Chinese firms.[35] Denoting this gap by τ^{TPU}, I estimate the following regression model using the OLS method:

$$log(w_{i,t+1}) = \beta_1 \tau_{jt}^{TPU} + \beta_2(\tau_{jt}^{TPU} \times s_{ijlt}) + \beta_3 s_{ijlt} + \gamma_i + \gamma_{lt} + \varepsilon_{it}. \tag{33}$$

In the regression model (33), $w_{i,t+1}$ is the observed (real) wage of firm i in year $(t+1)$.[36] τ_{jt}^{TPU} is the US-China TPU shock to industry j at year t, and is computed as:

$$\tau_{jt}^{TPU} = (\text{Column 2 Tariffs}_j - \text{MFN Tariffs}_j) \times PreWTO_t. \tag{34}$$

γ_i and γ_{lt} are firm and location-by-year fixed effects, respectively. s_{ijlt} is the local labor market wage-bill share of firm i, in industry j, within the location l, and in the year t. This wage-bill share is an empirical counterpart of the share defined in equation (15). Specifically, s_{ijlt} is computed as:

$$s_{ijlt} = \frac{w_{ijlt}L_{ijlt}}{\sum_i w_{ijlt}L_{ijlt}}. \tag{35}$$

Equation (33) is estimated with two outcome varibles: the (log) wage and the (log) employment.[37] I also estimate two versions of equation (33): one without β_2 and β_3, and the other with β_1, β_2, and β_3. The goal of this exercise is twofold. First, by comparing the average responses of wage and employment to the common τ_{jt}^{TPU} shock, I can identify the average labor supply elasticity of Chinese manufacturing firms. This gives me an estimate of the average endogenous distortion $\tilde{\chi}_i^e$ discussed in section 2.1. Second, by allowing for a firm's response to depend on its local labor market share, I can isolate the strategic component of labor market distortion. This approach is used by Berger, Herkenhoff and Mongey (2019) who study firms' labor market response to tax policy changes in the US, and more generally, in the trade literature to study the variable pass-through rate of international exchange rate shocks to firm-level prices, as in Amiti, Itskhoki and Konings (2019).

[35] "Column 2" tariffs are the tariffs assigned to nonmarket economies under the Smoot-Hawley Tariff Act of 1930. MFN tariffs are the tariffs offered to all members of WTO by the US.

[36] I use average wage, computed as the ratio of wage-bill and employment to measure the firm-level wage. Wage-bill data is deflated using detailed industry deflators at the four-digit industry level to account for any industry-specific trends, in the spirit of industry partial equilibrium models as in Melitz (2003). The industry deflators are obtained from Brandt et al. (2017).

[37] The regression equation for the (log) employment is:

$$log(L_{i,t+1}) = \beta_1 \tau_{jt}^{TPU} + \beta_2(\tau_{jt}^{TPU} \times s_{ijlt}) + \beta_3 s_{ijlt} + \gamma_i + \gamma_{lt} + \varepsilon_{it}. \tag{33A}$$

One could notice that the equations (33) and (33A) form a seemingly unrelated regression (SUR) system. Since the covariates are identical between the two equations, the OLS method would be equivalent to the SUR estimation method.

Table 3: Wage and Employment Response to the US-China TPU Demand Shock

Dependent variable	$log(w_{i,t+1})$		$log(L_{i,t+1})$	
	(1)	(2)	(3)	(4)
τ_{jt}^{TPU}	-0.075	-0.078	-0.118	-0.159
	(0.021)	(0.024)	(0.044)	(0.044)
Market Share (s_{ijlt}) $\times$ τ_{jt}^{TPU}		0.035		0.220
		(0.018)		(0.024)
Market Share (s_{ijlt})		0.011		0.213
		(0.005)		(0.007)
Observations	1,235,801	1,157,861	1,235,801	1,157,861
R-squared	0.711	0.734	0.897	0.911
Fixed Effects				
Firm	Yes	Yes	Yes	Yes
Location-Year	Yes	Yes	Yes	Yes
Clustered Two-way				
Firm	Yes	Yes	Yes	Yes
Industry-by-Year (4-digit)	Yes	Yes	Yes	Yes

Note: The table reports the results of regression equation (33) with two dependent variables: $log(w_{i,t+1})$ and $log(L_{i,t+1})$. Market Share (s_{ijlt}) is the local (prefecture) labor market wage-bill share, defined in equation (35), and τ_{jt}^{TPU} is the trade policy uncertainty (TPU) shock. Columns (2) and (4) have less observations because of the use of lagged share. Standard errors in parentheses are clustered two-way at firm level and industry-by-year level.

Identification of equation (33) is obtained by comparing changes in wage and employment of firms within the same location, yet exposed to differential labor demand shocks due to their industry affiliations. Any common time-varying local labor market shocks are controlled for by inclusion of location-by-year fixed effects. To allow time for firms to adjust their responses, and to alleviate a potential endogeneity concern of local labor market share, I use firms' outcomes in period $(t+1)$ as dependent variables to compute my estimates of labor supply elasticity. The coefficents of interest are β_1^h and β_2^h, where $h \in \{w, L\}$. Standard errors are clustered two-way, at firm-level and industry-by-year level, which is the variation level of the US-China TPU shock in these regressions.

Table 3 reports the regression results of equation (33). Columns (1) and (2) report results for equation (33) with the (log) wage as the dependent variable, and columns (3) and (4) report results for the same equation with the (log) employment as the dependent variable. Since the US-China TPU shock τ_{jt}^{TPU} is measured in log-form, the coefficients can be interpreted as percentage point changes. Let us first interpret the results from columns (1) and (3). Without the share-related covariates, results from these columns show that wage and employment both increased in response to the US-China TPU shock. Specifically, a one percentage point decrease in the "tariff uncertainty gap" τ_{jt}^{TPU} leads to a 0.075 point increase in wage and a 0.118 point increase in employment. On average, the change in τ_{jt}^{TPU} associated with China's accession to WTO is about 25 percentage points at the four-digit industry level, implying that the manufacturing wage and employment increased 1.88% and 2.95% respectively, due to the reduction in US-China TPU. More importantly, for the purpose of this study, these results suggest that the response of wage, i.e. $\frac{dlog(w)}{dlog(\tau^{TPU})}$, is about half the size of the response of the employment, i.e. $\frac{dlog(L)}{dlog(\tau^{TPU})}$. This result in turn indicates the *average* labor supply elasticity faced by a firm is 1.57. Computing the endogenous distortion from this labor supply elasticity implies the magnitude of endogenous distortion of 1.64. Compared with my average estimates from the production function approach in section 3, this endogenous distortion accounts for 76% of the overall distortion. This is one of the key findings of the paper.

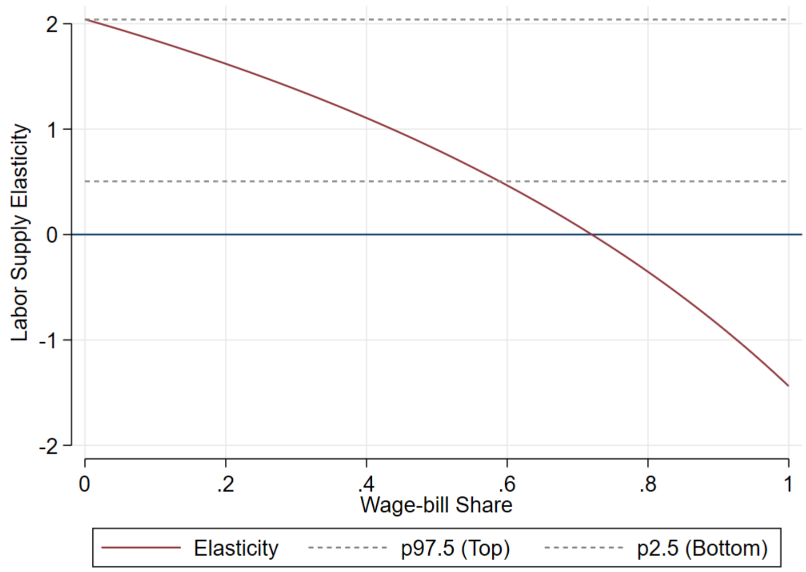
Estimation results in columns (2) and (4) further show that firms' response is nonlinear and varies by firms' local labor market share. This is reflected through the sign and magnitude of β_2^h. The interaction terms are positive and significant in both columns, suggesting that the response of firms with larger local labor market share is *weaker* to the US-China TPU demand shock.[38] To be more specific about the implication of these interaction coefficients, let us calculate the share-dependent labor supply elasticity based on the following formula:

$$\eta(s_{ijlt}) = \frac{dlog(L_{i,t+1})}{dlog(w_{i,t+1})} = \frac{\frac{dlog(L_{i,t+1})}{d(\tau_{jt}^{TPU})}}{\frac{dlog(w_{i,t+1})}{d(\tau_{jt}^{TPU})}} = \frac{\beta_1^L + \beta_2^L s_{ijlt}}{\beta_1^w + \beta_2^w s_{ijlt}}. \tag{36}$$

Given the formula in equation (36), it is useful to look at the labor supply elasticity for some particular values of s_{ijlt}. In Chinese firm-level data, the average local labor market share has decreased from 0.31 in 1998 to 0.17 in 2007. Plugging these numbers into equation (36), an average share of 0.31 implies the value of $\eta(0.31)$ is 1.35. An average share of 0.17 implies the value of $\eta(0.17)$ is 1.69. Firms with very small labor market share, 0.01 for instance, would face an elasticity of 2.01, while firms that account for a very large share of the local labor market, 0.5 for instance, would face a labor supply elasticity of 0.81. As a consequence, the endogenous distortion implied by the estimates in column (2) and (4) is much larger for firms that are the primary employers within a local labor market, i.e. these firms face a highly inelastic portion of the labor supply curve. To further illustrate the variation of labor supply elasticity as a function of local labor market share, Figure 4 graphs the computed

[38] This result resonates with the findings by Berger, Herkenhoff and Mongey (2019), in which the authors exploit changes in corporate taxes, rather than international trade shocks, to identify the endogenous distortion for the US.

Figure 4: Labor Supply Elasticity by Local Labor Market Wage-bill Share



Note: The figure illustrates the labor supply elasticity ($\eta(s_{ijlt})$) as a function of local labor market share in equation (36), with the estimated parameters obtained from regression equation (33). $p2.5$ and $p97.5$ are the 2.5th and 97.5th percentiles, with the values of 0.59 and 2.04 respectively.

elasticity based on equation (36).[39]

The results in Table 3 provide clear evidence that an endogenous form of labor market distortion exists. Such endogenous distortion accounts for almost 76% of the overall distortion measured by the production function approach in section 3. Furthermore, the distortion is significantly dependent on the local labor market share of firm. It is therefore possible to come up with a measure of the endogenous distortion that varies by firms' local labor market share.

A Measure of Share-Dependent Endogenous Distortion I use the share-dependent labor supply elasticity calculated by equation (36) to compute a measure of share-dependent endogenous distortion, the variation of which arises due to the actual variation of the local labor market share in the data. More specifically, from the share-dependent labor supply

[39]The estimated coefficients restrict the ability to infer the elasticity for firm with labor market share $s_{ijlt} \geq 0.723$ i.e. elasticity is negative. This implies that there might be further nonlinearity in the response of firms to the TPU shock, which I do not explore in this paper. However, there are less than 13% of firm-year observations that dominate the whole market, and I exclude these firms in my subsequent analyses when it involves the endogenous distortion measured by the regression approach in this section.

elasticity in equation (36), this measure of the endogenous distortion can be computed as:

$$\tilde{\chi}^e(s_{ijlt}) = 1 + \frac{1}{\eta(s_{ijlt})}. \tag{37}$$

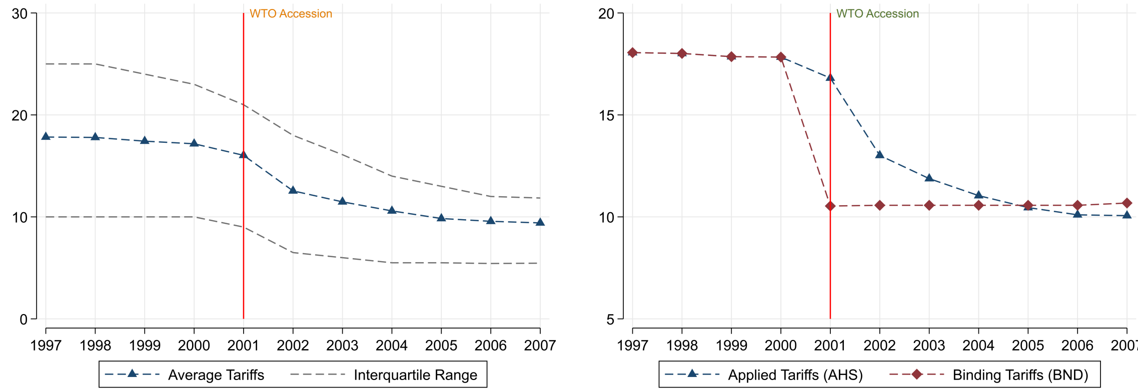
Measure of $\tilde{\chi}^e(s_{ijlt})$ as in equation (37) is consistent with the theory in section 2. However, it has a disadvantage: the only source of its variation comes from s_{ijlt}. In other words, measuring the endogenous distortion from the reduced-form approach as in equation (37) assumes that the structural parameters of the labor supply system in equations (6)-(7) are constant across industries and years. However, this measure is still useful and can serve two purposes: (1) it can help to cross-validate the measure of the distortion from the production function approach in section 3, and (2) it can illustrate how the response of local labor market shares to trade shocks translates directly to changes in labor market distortion.

Measure of distortion using parametric approach has quite limited variation as compared to the measure using production function approach. However, to cross-validate between the two measures, Figure D1-D2 reproduce Figure 2-3, respectively. The evolution and correlation patterns are very similar between the two measures. Specifically, in the right panel of Figure D1, the distribution of this endogenous distortion is becoming less dispersed over time and associated with reductions in both the mean and median. In terms of industry characteristics in Figure D2, industries that are more export-oriented and employ more female workers are also associated with lower level of labor market distortion. On the other hand, industries that have higher shares of state ownership and high-skill workers are associated with greater distortions. The only differences between the two measures are the *absolute* level of distortion and its across-firm dispersion, which is illustrated in the right panel of Figure D1. Since its only source of variation comes from the local labor market share, the distribution of the reduced-form measure is much less dispersed as compared to the production function measure. This resonates the results in Gaubert and Itskhoki (2021), in which they also find limited variation in product market distortion due to market shares. Nevertheless, in section 5, I compare the response of both measures to trade policy changes, and show that they exhibit remarkably similar patterns.

5 Impact of China's Trade Policy Reform

A key interest in this paper is to understand how China's own trade policy reform affected labor market distortion. In section 5.1, I briefly describe changes in China's trade policy regime associated with its accession to WTO in 2001. In section 5.2, I specify empirical models and present associated results.

Figure 5: China's Tariffs Evolution in the Pre-Post WTO Periods



Note: The left panel illustrates changes in the average and interquartile range of applied tariffs across industries and over time. The right panel illustrates changes in the (average) applied tariffs (AHS) versus binding tariffs (BND), which are used as instruments.

5.1 China's Trade Policy Regime upon WTO Accession

China's accession to WTO in December 2001 represents a major shift in China's trade policy regime over the past three decades. Upon accession, China committed to reduce the import tariffs from an average of 16 percent in the pre-WTO period to an average of 9 percent in the post-WTO period.[40] This paper focuses on the impact of lowering the tariff barriers. Specifically, I consider two policy instruments: the output tariff (τ^O) and the input tariff (τ^I), the empirical counterparts of the theoretical policy instruments in section 2. Input tariffs here are defined as input-share weighted averages of the output tariffs, using input expenditure shares from China's 2002 Input-Output table as weights, following Amiti and Konings (2007)'s approach in measuring input tariffs. In particular, the input tariff for industry j is calculated as:

$$\tau_j^I = \sum_m a_{mj} \tau_m^O, \tag{38}$$

where a_{mj} is the share of expenditure that industry j purchases from industry m, and τ_m^O is the output tariff that China imposes on industry m.

In the left panel of Figure 5, I plot the changes in the average and interquartile range of applied tariffs. In Table D3, I provide a summary of China's tariff evolution over the sample period from 1998-2007. The cutoff event is at the end of 2001, when China's official status in WTO became effective. As shown in Figure 5 and reported in Table D3, along

[40] Along with the reduction in import tariffs, China also made commiments to substantially reduce other non-tariff barriers upon WTO accession. See Brandt and Rawski (2008) and Brandt et al. (2017) for more institutional context of this event.

with a reduction in the average level of output tariffs, the standard deviation also fell from 9% in 1998 to 6% in 2007, implying that tariffs converged to a more uniform level across industries. A reduction in the standard deviation of tariffs across industries is the evidence for an exogenous source of tariff changes, and is commonly deployed in the empirical trade literature examining the effect of tariff liberalization on domestic outcomes (see for example, Amiti and Konings (2007), Topalova and Khandelwal (2011), De Loecker et al. (2016), Brandt et al. (2017)). Intuitively, a decrease in the dispersion of tariffs suggests that there is less room for the Home country's goverment to cherry-pick protection levels of specific industries due to political economy motives. For input tariffs, the tariff levels decreased from an average of 11% in 1998 to 6% in 2007, and the standard deviation decreased from 3% to 2% in respective years.[41]

5.2 Empirical Strategy

To investigate the causal impact of tariff liberalization on the endogenous response of firm-level distortion, I adopt a version of the empirical specifications widely used in the empirical trade literature, for examples, as in Pavcnik (2002), Amiti and Konings (2007), Topalova and Khandelwal (2011), and Brandt et al. (2017). The specification is as follows:

$$log(\tilde{\chi}_{ijlt}) = \gamma_O \times \tau^O_{j,t-1} + \gamma_I \times \tau^I_{j,t-1} + \gamma_i + \gamma_{cic2,t} + \gamma_{lt} + \varepsilon_{ijlt}. \tag{39}$$

In equation (39), the dependent variable is the (log) measured labor market distortion. $\tau^O_{j,t-1}$ and $\tau^I_{j,t-1}$ are the one-year *lagged* output and input tariffs for industry j, computed at four-digit and three-digit aggregation level respectively. γ_i controls for firms' fixed effects, and γ_{lt} controls for location-by-year fixed effects, similar to the regression in equation (33). Since my production function is estimated at the two-digit industry level, I supplement my analysis with $\gamma_{cic2,t}$, which controls for any time-varying changes at the two-digit industry level that may confound the results.[42] The coefficients of interest are γ_O and γ_I.[43] Intuitively, these coefficients are identified by comparing the differential *changes* of the outcome variable across firms, within the same location-by-year and cic2-by-year group. These firms differ only in

[41] Due to the aggregation level of China's 2002 Input-Output table, input tariffs only vary at the three-digit industry level, which contributes to the lesser degree of variation in the input tariffs across industries in Table D3. The correlation between output and input tariff changes across 4-digit industries is 0.59, and including them either individually or simultaneously almost do not change the estimates.

[42] This additional fixed-effect term turns out to be important, since I observe some diverging trends in measured distortions and productivity at the 2-digit industry level.

[43] An essential assumption for identification of the causal impact of tariff changes on the labor market distortion using equation (39) is the constant treatment effect assumption, in which I assume that the causal effects of tariffs are constant across firm, location, industry and time.

their differential exposure to *changes* in tariffs at four-digit (or three-digit) industry level. Across all the specifications, standard errors are clustered two-way, at firm and industry-by-year level.

In my baseline estimation, equation (39) is estimated with the OLS method. Although there is a large set of fixed-effects included in equation (39), there still might be an endogeneity concern of the industry-level tariff changes. Recall that, in my theoretical analysis, productivity is a key determinant of firm-level labor market distortion. When China joined the WTO in 2001, it is possible that the Chinese government selectively reduced tariffs for certain industries based on their past productivity growth trends. If this is indeed the case, differences in labor market distortion of firms across industry might be attributable to industry-specfic growth in productivity rather than being caused by differential tariff changes. I address this endogeneity concern by following the identification approach in Brandt et al. (2017). Specifically, for the post-WTO period (after 2001), I use the maximum binding tariff negotiated (and fixed) in 1999 as an instrumental variable for the actual applied tariff, and estimate equation (39) with the 2SLS method. The right panel of Figure 5 illustrates the co-evolution of the average applied and binding tariff over time. This instrument alleviates the policy endogeneity concern because it is presumably difficult for Chinese policymakers to correctly predict the productivity evolution of various industries in the post-WTO period and negotiate the maximum tariff levels accordingly. In what follows, my preferred quantitative interpretation is based on the IV estimates.

Entry, Exit, and Labor Market Share To begin with, Table 4 reports the estimation results with three dependent variables: (1) indicator for firm entry, (2) indicator for firm exit, and (3) log of labor market share ($\log(s_{ijlt})$), defined in section 4. Columns (1)-(4) show that entry and exit patterns across local labor markets respond significantly to tariff changes. In particular, lowering output tariffs decreases probability of a firm entering and increases probability in which a firm exiting the local market. On the other hand, lowering input tariffs increases probability of entry while reducing probability of exit.[44] These results resonate the findings in previous literature where plant survival rate, growth, and consequential labor market outcomes are associated with trade shocks (Bernard, Jensen and Schott (2006), Asquith et al. (2019)). Labor market share also responds to trade policy changes. Specifically, in column (5)-(6), lowering output tariffs is negatively associated with labor market share

[44]It is nevertheless important to emphasize that the definition of entry and exit here depends on the sample. That is, one should interpret these results as entering and exiting the working sample, which in this case involves all firms with sales above 5 millions Renminbi ($\approx$ 700000 US Dollars in current price). My cleaning procedure keeps the size threshold consistent and these results remain very robust to less stringent cleaning procedures.

Table 4: Impact of Tariff Changes on Entry, Exit, and Labor Market Share

Dependent Variable	Entry		Exit		Market Share	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
Output Tariffs ($\tau^O_{j,t-1}$)	0.110	0.084	-0.094	-0.078	-0.124	-0.210
	(0.037)	(0.042)	(0.027)	(0.029)	(0.199)	(0.235)
Input Tariffs ($\tau^I_{j,t-1}$)	-0.788	-0.894	0.341	0.291	4.656	6.304
	(0.255)	(0.332)	(0.131)	(0.141)	(2.547)	(3.251)
Observations	1,235,801	1,235,801	1,235,801	1,235,801	1,235,801	1,235,801
R-squared	0.524	.	0.742	.	0.927	.
Fixed Effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Location-Year	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Two-way						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-Year (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports the results of regression equation (39) with three dependent variables: (1) indicator for firm entry, (2) indicator for firm exit, and (3) log of labor market share ($\log(s_{ijlt})$), defined in section 4. All tariffs are measured as the natural log of 1 plus the ad valorem tariffs i.e. $ln(1+\tau)$. Instruments for the applied tariffs (after 2001) are the maximum tariffs negotiated before China's accession to WTO. Standard errors in parentheses are clustered two-way at firm and industry-by-year level.

of a firm, while lowering input tariffs is positively associated with this firm-level variable. However, the estimated coefficients are only significant for input tariff reductions, implying a stronger influence of input tariffs in affecting labor market share. Overall, Table 4's results are strongly consistent with the prediction of theory, which lends empirical supports for the modeling approach proposed in section 2.2.

Labor Market Distortion Table 5 reports the estimation results for this paper's main outcomes of interest. Columns (1)-(4) show the results of the regression equation (39), using the (log) overall distortion measured in section 3 as the dependent variable. Across the columns, which use two complementary estimation methods, i.e. OLS and IV, the sign and the significance of the coefficients estimated are consistent with the predictions of theory in propositions (2) and (3). The coefficients of the output tariff has a negative sign, while the input tariff coefficient has a positive sign. In other words, the results show that a reduction in the output tariff increases the labor market distortion, while a reduction in the input tariff leads to a decrease in the measured distortion.

Table 5: Impact of Tariff Changes on the Labor Market Distortion

	Nonparametric Measure				Parametric Measure			
Dependent Variable	$log(\tilde{\chi}_{ijlt})$				$log(\tilde{\chi}^e(s_{ijlt}))$			
	OLS (1)	OLS (2)	OLS (3)	IV (4)	OLS (5)	OLS (6)	OLS (7)	IV (8)
Output Tariffs ($\tau^O_{j,t-1}$)	-0.012		-0.038	-0.029	-0.024		-0.035	-0.035
	(0.058)		(0.057)	(0.053)	(0.021)		(0.020)	(0.021)
Input Tariffs ($\tau^I_{j,t-1}$)		0.484	0.531	0.592		0.220	0.259	0.327
		(0.267)	(0.260)	(0.291)		(0.140)	(0.140)	(0.160)
Observations	1,235,801	1,235,801	1,235,801	1,235,801	1,040,013	1,040,013	1,040,013	1,040,013
R-squared	0.758	0.758	0.758	.	0.668	0.668	0.668	.
Fixed Effects								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location-Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Two-way								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-Year (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports the results of regression equation (39) with two dependent variables: $log(\tilde{\chi}_{ijlt})$ and $log(\tilde{\chi}^e(s_{ijlt}))$. All tariffs are measured as the natural log of 1 plus the ad valorem tariffs i.e. $ln(1+\tau)$. Instruments for the applied tariffs (after 2001) are the maximum tariffs negotiated before China's accession to WTO. Obsevations with negative values of the labor supply elasticity measured in section 4 are trimmed in columns (5)-(8). Standard errors in parentheses are clustered two-way at firm and industry-by-year level.

Quantitatively, however, only input tariffs have strong and significant impact on the distortion. Specifically, a one percentage point decrease in input tariffs leads to a 0.53 percentage point reduction in the distortion, based on the OLS estimates. IV estimates imply a 0.59 percentage point decrease in the distortion. Combining with the actual change in the input tariffs during my sample period, in which the input tariffs decreased from 11% to 6% on average from 1998-2007, these estimation results suggest that the average labor market distortion decreased by 2.95% as a consequence of China's reduction in the input tariffs. Compared with the effect of input tariffs, output tariff reductions had negligible and insignificant effects on labor market distortion.

Columns (5)-(8) estimate equation (39) using the (log) share-dependent distortion measured by the regression (parametric) approach in section 4 as the dependent variable. The goal of these columns is to compare the results between the two different measures of labor market distortion. Consistent with the results in columns (1)-(4), estimated coefficients for tariffs have the same signs. In columns (7)-(8), however, the coefficients of both tariffs are significant, with the reduction in input tariffs having a much stronger effect. Across the two columns, a one percentage point decrease in output tariffs leads to a 0.035 percentage point increase in the share-dependent distortion in the OLS estimates, compared to the same estimated of 0.035 in the IV estimates. On the other hand, a one percentage point decrease in the input tariffs leads to a 0.259 and 0.327 percentage point decrease in the share-dependent distortion respectively in the OLS and IV estimates. Taken together and combined with actual tariff changes during the sample period 1998-2007, columns (7) and (8) suggest that the output tariff reduction led to a 0.25% increase in the distortion, while the input tariff reduction led to a 1.96% decrease in the distortion (using IV estimates). Remarkably, these estimates are very similar and in the same order of magnitude as compared to the production function measure of the distortion, despite employing two totally different approaches to measurement. The results imply that there is a common underlying economic mechanism driving the responses of measured distortions to trade policy changes, which lends further supports to the presence of competitive effects of trade in the labor market.

Markup It is theoretically possible that changes in labor market distortions in response to trade policy are results of changes in product market distortions. In MacKenzie (2019), trade shocks can induce reallocation in an oligopolistic product market, and thus induce reallocation in labor market. This is due to the fact that firms with larger product market share also tend to be the ones with larger labor market share (i.e. more productive firms). To investigate this mechanism, I exploit my production function estimation in section 3.1 to measure firm-level markups. Here, due to the lack of firm-level price information, I follow

Table 6: Impact of Tariff Changes on Markup

Dependent Variable	Markup					
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Output Tariffs ($\tau^{O}_{j,t-1}$)	-0.006	-0.008			-0.004	-0.006
	(0.003)	(0.003)			(0.003)	(0.003)
Input Tariffs ($\tau^{I}_{j,t-1}$)			-0.041	-0.046	-0.035	-0.037
			(0.015)	(0.019)	(0.017)	(0.020)
Observations	1,235,801	1,235,801	1,235,801	1,235,801	1,235,801	1,235,801
R-squared	0.873	.	0.873	.	0.873	.
Fixed Effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Location-Year	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Two-way						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-Year (2-digit)	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports the results of regression equation (39) with the dependent variable as the log of markup. All tariffs are measured as the natural log of 1 plus the ad valorem tariffs i.e. $ln(1+\tau)$. Instruments for the applied tariffs (after 2001) are the maximum tariffs negotiated before China's accession to WTO. Standard errors in parentheses are clustered two-way, at firm and industry-by-year level.

Flynn, Gandhi and Traina (2019) in imposing the constant return-to-scale assumption to infer the markup and regress this measure on China's trade policy changes as in equation (39). The results are shown in Table 6.

Across columns (1)-(6), a robust result is that lowering input tariffs increases firm-level markups. This result is statistically significant and qualitatively consistent with previous findings in the literature (De Loecker et al. (2016), Brandt et al. (2017)). On the other hand, lowering output tariffs almost has no substantive economic impacts on firm-level markups.[45] Comparing the magnitude of the effect on markups versus the effect on labor market distortion in table 5, the effect of trade policy on markups is an order of magnitude smaller. This implies that changes in markups cannot explain the changes in labor market distortions in response to China's tariff changes.

Summing Up Overall, the empirical results in this section confirm predictions in propositions (2) and (3) about the impact of a trade reform on firm-level labor market distortion.

[45]In anything, lowering output tariffs actually increases the markup slightly, implying that product market power increases due to competitive output shocks. The result that lowering output tariffs increases markups is consistent with an oligopoly model in which competitive output shocks kill off less productive firms and reallocate product market share towards more productive, surviving firms, similar in spirit to MacKenzie (2019).

My prefered estimates suggest that even though lowering output tariffs has a tendency to increase the distortion, its effect is small. On the other hand, lowering input tariffs substantially reduces the labor market distortion, with the magnitude of the overall effect of 2.95%. These results hold robustly when using alternative measures of labor market distortion and cannot be explained by changes in product market distortion (markup).

6 Conclusion

This paper studies the impact of international trade policy on competition in the labor market. The paper makes three contributions. First, I develop a tractable model to study the impact of trade policy on distortions in the labor market, providing clear predictions based on this model. Second, I propose two complementary strategies to consistently measure labor market distortion and show that the magnitude of this distortion can be large, contradicting a critical assumption in many trade models that the labor market is perfectly competitive. Third, I establish a causal relationship between trade policy and the endogenous labor market distortion. A key takeaway is that opening up to import competition through lowering output tariffs potentially increases the distortion in the labor markets. On the other hand, lowering input tariffs can substantially decrease the distortion by allowing firms to access cheaper foreign inputs. The proposed operating mechanism of such effects is the endogenous entry and exit of firms across local labor markets induced by trade shocks.

My theoretical and empirical results have a number of implications for our understanding of how trade policy affects labor market performance. Since labor market power has consequential effects on wages, employment, labor shares and inequality, my results suggest that trade can affect the labor market power of firms and thus, alters the labor market outcomes through this mechanism. Even though the context of my empirical analysis is a developing country, i.e. China, it is plausible that this mechanism also operates in developed economies. A fruitful direction for future research is thus to analyze my results' generalizability to a developed country context. Furthermore, as endogenous distortion accounts for a major part of overall labor market distortion, my results suggest that standard welfare calculations of trade, notably as in Arkolakis, Costinot and Rodriguez-Clare (2012), might be affected by the presence of such distortion and its endogenous response to trade.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83: 2411–2451.

Amiti, Mary, and Donald R Davis. 2011. "Trade, Firms, and Wages: Theory and Evidence." *The Review of Economic Studies*, 79(1): 1–36.

Amiti, Mary, and Jozef Konings. 2007. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review*, 97(5): 1611–1638.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International Shocks, Variable Markups, and Domestic Prices." *The Review of Economic Studies*, 86(6): 2356–2402.

Anderson, Simon P, André De Palma, and Jacques-Francois Thisse. 1987. "The CES Is a Discrete Choice Model?" *Economics Letters*, 24(2): 139–140.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review*, 102(1): 94–130.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2018. "The Elusive Pro-competitive Effects of Trade." *The Review of Economic Studies*, 86(1): 46–80.

Ashenfelter, Orley C, Henry Farber, and Michael R Ransom. 2010. "Labor Market Monopsony." *Journal of Labor Economics*, 28(2): 203–210.

Asquith, Brian, Sanjana Goswami, David Neumark, and Antonio Rodriguez-Lopez. 2019. "US Job Flows and the China Shock." *Journal of International Economics*, 118: 123–137.

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review*, 98(5): 1998–2031.

Benmelech, Efraim, Nittai K Bergman, and Hyunseob Kim. 2020. "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?" *Journal of Human Resources*, 0119–10007R1.

Berger, David W, Kyle F Herkenhoff, and Simon Mongey. 2019. "Labor Market Power." National Bureau of Economic Research Working Paper 25719.

Bernard, Andrew B, J Bradford Jensen, and Peter K Schott. 2006. "Survival of the Best Fit: Exposure to Low-wage Countries and the (Uneven) Growth of US Manufacturing Plants." *Journal of international Economics*, 68(1): 219–237.

Bernard, Andrew B, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *American Economic Review*, 93(4): 1268–1290.

Boal, William M., and Michael R Ransom. 1997. "Monopsony in the Labor Market." *Journal of Economic Literature*, XXXV: 86–112.

Brandt, Loren, and Thomas G Rawski. 2008. *China's Great Economic Transformation.* Cambridge University Press.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang. 2017. "WTO Accession and Performance of Chinese Manufacturing Firms." *American Economic Review*, 107: 2784–2820.

Brandt, Loren, Johannes Von Biesebroeck, and Yifan Zhang. 2014. "Challenges of Working with the Chinese NBS Firm-level Data." *China Economic Review*, 30: 339–352.

Brooks, Wyatt J, Joseph P Kaboski, Yao Amber Li, and Wei Qian. 2021. "Exploitation of Labor? Classical Monopsony Power and Labor's Share." *Journal of Development Economics*, 150: 102627.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica*, 79(5): 1407–1451.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-level Export Status." *American Economic Review*, 102(6): 2437–71.

De Loecker, Jan, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform." *Econometrica*, 84(2): 445–510.

Dobbelaere, Sabien, and Quint Wiersma. 2019. "The Impact of WTO Accession on Chinese Firms' Product and Labor Market Power." Tinbergen Institute Discussion Paper 2019-037/V.

Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. 2020. "Monopsony in Online Labor Markets." *American Economic Review: Insights*, 2(1): 33–46.

Eaton, Jonathan, Samuel S Kortum, and Sebastian Sotelo. 2012. "International Trade: Linking Micro and Macro." National Bureau of Economic Research Working Paper 17864.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, Markups, and the Gains from International Trade." *American Economic Review*, 105(10): 3183–3221.

Flynn, Zach, Amit Gandhi, and James Traina. 2019. "Measuring Markups with Production Data." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3358472. Accessed on 2021-03-02.

Gandhi, Amit, Salvador Navarro, and David A Rivers. 2020. "On the Identification of Gross Output Production Functions." *Journal of Political Economy*, 128(8): 2973–3016.

Gaubert, Cecile, and Oleg Itskhoki. 2021. "Granular Comparative Advantage." *Journal of Political Economy*, 129(3): 000–000.

Goldberg, P.K., and N. Pavcnik. 2016. "The Effects of Trade Policy." *Handbook of Commercial Policy*, 1A: 161–206.

Halpern, László, Miklós Koren, and Adam Szeidl. 2015. "Imported Inputs and Productivity." *American Economic Review*, 105(12): 3660–3703.

Handley, Kyle, and Nuno Limao. 2015. "Trade and Investment under Policy Uncertainty: Theory and Firm Evidence." *American Economic Journal: Economic Policy*, 7(4): 189–222.

Handley, Kyle, and Nuno Limão. 2017. "Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States." *American Economic Review*, 107(9): 2731–83.

Hau, Harald, Yi Huang, and Gewei Wang. 2020. "Firm Response to Competitive Shocks: Evidence from China's Minimum Wage Policy." *The Review of Economic Studies*, 87(6): 2639–2671.

Heiland, Inga, and Wilhelm Kohler. 2019. "Heterogeneous Workers, Trade, and Migration."

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. 2010. "Inequality and Unemployment in a Global Economy." *Econometrica*, 78: 1239–1283.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Hsieh, Chang-Tai, and Zheng Michael Song. 2015. "Grasp the Large, Let Go of the Small:: The Transformation of the State Sector in China." *Brookings Papers on Economic Activity*, , (1): 3.

Kucheryavyy, Konstantin. 2012. "Continuity of A Model with A Nested CES Utility Function and Bertrand Competition." *Economics Letters*, 117(2): 473–476.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70: 317–341.

Liu, Ernest. 2019. "Industrial Policies in Production Networks." *The Quarterly Journal of Economics*, 134(4): 1883–1948.

Lu, Yi, and Linhui Yu. 2015. "Trade Liberalization and Markup Dispersion: Evidence from Chinas WTO Accession." *American Economic Journal: Applied Economics*, 7(4): 221253.

Lu, Yi, Yoichi Sugita, and Lianming Zhu. 2019. "Wage and Markdowns and FDI Liberalization." Hitotsubashi Institute for Advanced Study, Hitotsubashi University Discussion Paper Series HIAS-E-83.

MacKenzie, W. Gaelan. 2019. "Trade and Market Power in Product and Labor Markets." https://www.dropbox.com/s/ml3j3tg3zzenmsc/JMP_Latest.pdf?dl=0. Accessed on 2021-03-02.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71: 1695–1725.

Melitz, Marc J, and Gianmarco I.P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *The Review of Economic Studies*, 75(1): 295–316.

Morlacco, Monica. 2020. "Market Power in Input Markets: Theory and Evidence from French Manufacturing." https://www.dropbox.com/s/4ic4ylvgngzn947/BP_draft_01102020.pdf?dl=0. Accessed on 2021-03-02.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64: 1263–1297.

Pavcnik, Nina. 2002. "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants." *The Review of Economic Studies*, 69(1): 245–276.

Pierce, Justin R, and Peter K Schott. 2016. "The Surprisingly Swift Decline of US Manufacturing Employment." *American Economic Review*, 106(7): 1632–62.

Pigou, Arthur. 1924. *The Economics of Welfare.* Routledge.

Ranjan, Priya, and Antonio Rodriguez-Lopez. 2019. "Monopsonistic Labor Markets and International Trade." CESifo Working Paper Series 7771.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic dynamics*, 11(4): 707–720.

Robinson, Joan. 1969. *The Economics of Imperfect Competition.* Springer.

Staiger, Douglas O, Joanne Spetz, and Ciaran S Phibbs. 2010. "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment." *Journal of Labor Economics*, 28(2): 211–236.

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review*, 109(5): 1843–72.

Topalova, Petia, and Amit Khandelwal. 2011. "Trade Liberalization and Firm Productivity: The Case of India." *Review of Economics and Statistics*, 93(3): 995–1009.

Verboven, Frank. 1996. "The Nested Logit Model and Representative Consumer Theory." *Economics Letters*, 50(1): 57–63.

For Online Publication

A. Theory Appendix

A.1 Key model results

First-order conditions (FOC) First-order condition of firm's profit maximization problem:

$$\frac{MRPL(z)}{w(z)} = \frac{\frac{\gamma-1}{\gamma} z p(z)}{w(z)} = 1 + \frac{1}{\varepsilon(s(z))}, \tag{A1}$$

where $\varepsilon(s(z))$ is the elasticity of labor supply and has value:[46]

$$\varepsilon(s(z)) = \frac{1}{\frac{1}{\eta} + s(z)(\frac{1}{\theta} - \frac{1}{\eta})}. \tag{A2}$$

Equilibirum wages We can also rewrite $MRPL(z)$ as:

$$\begin{aligned} MRPL(z) &= z^{\frac{\gamma-1}{\gamma}} L(z)^{-\frac{1}{\gamma}} \mathbf{P}^{\frac{\gamma-1}{\gamma}} \mathbf{\Xi} \\ &= z^{\frac{\gamma-1}{\gamma}} w(z)^{-\frac{\theta}{\gamma}} s(z)^{-\frac{\Psi}{\gamma}} \mathbf{\Lambda} \end{aligned} \tag{A3}$$

where $\mathbf{\Lambda} = \bar{\phi}^{-\frac{1}{\gamma}} \mathbf{W}^{\frac{\theta-\phi}{\gamma}} \mathbf{P}^{\frac{\gamma-1}{\gamma}} \mathbf{\Xi} > 0$ and $\Psi = \frac{\eta-\theta}{1+\eta} > 0$.[47] Combining (A1) and (A3), we can derive the following equilibrium relationship:

$$\begin{aligned} w(z) &= \left(\frac{1}{[(1+\frac{1}{\eta}) + (\frac{1}{\theta} - \frac{1}{\eta}) s(z)] s(z)^{\frac{\Psi}{\gamma}}} \right)^{\frac{\gamma}{\gamma+\theta}} z^{\frac{\gamma-1}{\gamma+\theta}} \mathbf{\Lambda}^{\frac{\gamma}{\gamma+\theta}} \\ &= m(s(z)) \ddot{z}. \end{aligned} \tag{A4}$$

Here, we denote $m(s(z)) = \left(\frac{1}{[(1+\frac{1}{\eta}) + (\frac{1}{\theta} - \frac{1}{\eta}) s(z)] s(z)^{\frac{\Psi}{\gamma}}} \right)^{\frac{\gamma}{\gamma+\theta}}$ and $\ddot{z} = z^{\frac{\gamma-1}{\gamma+\theta}} \mathbf{\Lambda}^{\frac{\gamma}{\gamma+\theta}}$. It is straightforward to verify that $m(s(z))$ is a *strictly decreasing* function in $s(z)$, while $\ddot{z}$ is a *strictly increasing* function in z and $\mathbf{\Lambda}$ (since $\theta < \eta$, $\Psi > 0$, and $\gamma > 1$). This equation illustrates a key property of the model that makes the following proofs work. Finally, FOCs of all firms

[46]The expression for $\varepsilon(s(z))$ equation can be derived straightforwardly from the labor supply curve in equation (11).

[47]This can be shown by noticing that $L(z) = \bar{\phi} w(z)^{\eta} \mathbf{W}_n^{\theta-\eta} \mathbf{W}^{\phi-\theta} = \bar{\phi} \frac{w(z)^{\eta}}{w(z)^{\eta-\theta}} \frac{w(z)^{\eta-\theta}}{\left[\sum_{z' \in Z_n} w(z')^{1+\eta}\right]^{\frac{\eta-\theta}{1+\eta}}} \mathbf{W}^{\phi-\theta} = \bar{\phi} w(z)^{\theta} s(z)^{\frac{\eta-\theta}{1+\eta}} \mathbf{W}^{\phi-\theta}$.

form the following system of equations in the vector form:

$$\bar{w}(z) = \bar{m}(s(z))^T \bar{\bar{z}} \tag{A5}$$

A.2 Profit monotonicity

The *profit monotonicity* condition is stated as follows:

$$\Pi^K(z^{(k)}) \geq \Pi^{K+1}(z^{(k)}) \geq \Pi^{K+1}(z^{(k+1)}). \tag{A6}$$

The second inequality is proved first. This inequality states that, given the same market condition, more productive firms always command higher profits than less productive firms (recall the productivity ranking $z^{(1)} > ... > z^{(k)} > z^{(k+1)}$). The proof is straightforward by deduction: given the aggregate equilibrium conditions, the more productive firm $z^{(k)}$ can always hire the same amount of workers, pay the same wages, charge at least the same price as the less productive firm $z^{(k+1)}$, and makes more profit (see also in Eaton, Kortum and Sotelo (2012)). (Q.E.D)

The first inequality states that for the same firm with productivity $z^{(k)}$, the firm is more profitable if there are less competitors in the local labor market, i.e. removing the firm $z^{(k+1)}$ from the market. To show this, let's first rewrite the equilibrium profit of firm z as:

$$\begin{aligned} \Pi(z) &= \Big[\frac{\gamma}{\gamma-1}(1+\frac{1}{\varepsilon(s(z))}) - 1\Big] w(z)L(z) - f \\ &= w(z)^{\theta+1}\Big[\frac{\gamma}{\gamma-1}\Big(1+\frac{1}{\eta}\Big) + s(z)\Big(\frac{1}{\theta} - \frac{1}{\eta}\Big) - 1\Big] s(z)^{\frac{\eta-\theta}{1+\eta}} \bar{\phi} \mathbf{W}^{\phi-\theta} - f \end{aligned} \tag{A7}$$

Base on this profit function, it can be shown that, for $z' \neq z$:[48]

$$\frac{\partial \Pi(z)}{\partial w(z')} \leq 0 \qquad \text{and} \qquad \frac{\partial^2 \Pi(z)}{\partial w(z) \partial w(z')} \geq 0. \tag{A8}$$

The left inequality asserts that lower wage on the part of a local competitor increases profit of firm z. This is because lower $w(z')$ reduces the employment of firm z', and thus allow firm z to set lower wage to attract workers. The right inequality asserts that lower wage of a local competitor induces firm z to decrease its wage (wage complementarity).

From inequalities in (A8), by removing firm $z^{(k+1)}$ and essentially makes its wage $w(z^{(k+1)}) \to 0$, all incumbent firms in the market becomes more profitable, and hence, $\Pi^K(z^{(k)}) \geq \Pi^{K+1}(z^{(k)})$. (Q.E.D)

[48] See also similar conditions and arguments for the oligopoly model in Eaton, Kortum and Sotelo (2012). Derivations for these inequalities are available upon requests.

A.3 Proposition 1

Proposition 1 is proved in two steps. First, we show that conditioning on a fixed number of firm K, there exists a unique equilibrium vector of wages (and associated vectors of employment and shares) in the local labor market.

Existence Expression in equation (A4) is sometimes sufficient to bound the wages and allows us to invoke Brouwer's fixed point theorem directly for existence. However, in this case, due to the presence of CES demand ($\gamma < \infty$), wages cannot be bounded as $s(z) \to 0$. It is thus useful to rewrite the system of equations in shares. Equations in (A4) and (A5) are equivalent to:

$$w(z)^{\eta+1} = m(s(z))^{\eta+1}\ddot{z}^{\eta+1} \tag{A9}$$
$$\dots$$

Sum up these equations vertically across all firms, we have:

$$\sum_z w(z)^{\eta+1} = \sum_z m(s(z))^{\eta+1}\ddot{z}^{\eta+1} \tag{A10}$$

Divide each equation in (A9) to (A10) and using the share expression in equation (15), we obtain:[49]

$$s(z) = \frac{m(s(z))^{\eta+1}\ddot{z}^{\eta+1}}{\sum_z m(s(z))^{\eta+1}\ddot{z}^{\eta+1}} \tag{A11}$$

From this equation, we can derive the lower bound of $s(z)$. Let $s_{min} = s(z^{(K)})$ such that $s_{min} \leq s(z)\ \forall z$. Recall that since $m(.)$ is strictly decreasing in $s(z)$ and $\eta > 0$, we must have:

$$s_{min} = \frac{m(s_{min})^{\eta+1}\ddot{z}_{min}^{\eta+1}}{\sum_z m(s(z))^{\eta+1}\ddot{z}^{\eta+1}} \geq \frac{m(s_{min})^{\eta+1}\ddot{z}_{min}^{\eta+1}}{m(s_{min})^{\eta+1}\left[\sum_z \ddot{z}^{\eta+1}\right]} = \frac{\ddot{z}_{min}^{\eta+1}}{\sum_z \ddot{z}^{\eta+1}} = s^* > 0 \tag{A12}$$

As a result, s^* is the lower bound of labor market share.

Define a set S as: $\bar{s} \in [s^*, 1]^K$ and $\sum_{z \in Z_K} s(z) = 1$. By construction, S is a nonempty, compact and convex set. Define a function $F : S \mapsto S$ as follow:

$$F_z(\bar{s}) = \frac{m(s(z))^{\eta+1}\ddot{z}^{\eta+1}}{\sum_z m(s(z))^{\eta+1}\ddot{z}^{\eta+1}} \qquad \forall z \in Z_K \tag{A13}$$

[49] Notice that this equation is free of aggregate variables contained in $\mathbf{\Lambda}$ and vector of equilibrium shares only depend the set of firms' productivity, since $\mathbf{\Lambda}$ cancels out.

Function $F : S \mapsto S$ is a continuous function mapping from S to itself. Thus, Brouwer's fixed point theorem applies and the existence result is established.(Q.E.D)

Uniqueness Suppose that there exists two different vectors of wages satisfying system of equations (A5). As a result, they are characterized by two different vectors of shares $(s'_1, ..., s'_K)$ and $(s''_1, ..., s''_K)$. Hereafter, productivity ranking of firms will not matter for the proof, thus I index firms by subscript k rather than productivity z. It follows that $s'_\kappa \neq s''_\kappa$ for some $1 \leq \kappa \leq K$.

The following proof mimics closely the proof proposed by Kucheryavyy (2012) for an oligopolistic competition model.[50] Without loss of generality, suppose that $s'_\kappa > s''_\kappa$. It follows from equation (A4) that $w'_\kappa < w''_\kappa$ since $m'(s_\kappa) < 0$.[51] It is also follows that $\left(\frac{w'_\kappa}{w''_\kappa}\right)^{\eta+1} < 1$.

For any $k = 1, \ldots, K$, let $r_k = \left(\frac{w'_k}{w''_k}\right)^{\eta+1}$. We reorder firms by index k such that $r_k \leq r_l$ for any $k \leq l$. Since $r_\kappa < 1$ for some κ, we must have $r_1 < 1$. For any k, denote:

$$\begin{aligned} d'_k &\equiv (w'_k)^{\eta+1} + \ldots + (w'_K)^{\eta+1} \\ d''_k &\equiv (w''_k)^{\eta+1} + \ldots + (w''_K)^{\eta+1} \end{aligned} \tag{A14}$$

We will show that $\frac{d'_k}{d''_k} < r_1 \ \forall k \geq 2$.

First, consider the case where $k = 2$. By definition and from the share expression in equation (15), $r_1 < 1$ is equivalent to $s'_1 > s''_1$ and thus:

$$\begin{aligned} &\frac{(w'_1)^{\eta+1}}{(w'_1)^{\eta+1} + d'_2} > \frac{(w''_1)^{\eta+1}}{(w''_1)^{\eta+1} + d''_2} \\ \iff &\frac{d'_2}{d''_2} < \left(\frac{w'_1}{w''_1}\right)^{\eta+1} = r_1 \end{aligned} \tag{A15}$$

Next, suppose that $\frac{d'_k}{d''_k} < r_1$ for some $k \geq 2$, we will show that $\frac{d'_{k+1}}{d''_{k+1}} < r_1$. By construction,

[50] The proof by Kucheryavyy is provided on his website. I am extremely grateful to Oleg Itskhoki and Konstantin Kucheryavyy for refering me to this proof.

[51] An attentive reader will notice that $\ddot{z}$ is a function of $\mathbf{W}$, which in turn is function of s_κ itself. However, recall that because there is a continuum of local labor markets, vector of shares $\bar{s_k}$ within a local labor market has no influence on the aggregate wage index and thus, firms take this term as given. Alternatively, one can make this proof works by directly using decreasing property of function $m(.)$ in system of equations (A11) and (A13) since the influence of aggregate variables in $\mathbf{\Lambda}$ cancels out .

$\frac{d'_k}{d''_k} < r_1$ is equivalent to:

$$\begin{aligned} &\frac{(w'_k)^{\eta+1} + d'_{k+1}}{(w''_k)^{\eta+1} + d''_{k+1}} < r_1 \\ \iff &(w'_k)^{\eta+1} + d'_{k+1} < r_1\Big((w''_k)^{\eta+1} + d''_{k+1}\Big) \end{aligned} \tag{A16}$$

Divide both sides by d'_{k+1}, we have:

$$\begin{aligned} &\frac{(w'_k)^{\eta+1}}{d'_{k+1}} + 1 - r_1\frac{(w''_k)^{\eta+1}}{d'_{k+1}} < r_1\Big(\frac{d''_{k+1}}{d'_{k+1}}\Big) \\ \iff 1 + &\frac{(w'_k)^{\eta+1}}{d'_{k+1}}\Big(1 - r_1\frac{(w''_k)^{\eta+1}}{(w'_k)^{\eta+1}}\Big) < r_1\Big(\frac{d''_{k+1}}{d'_{k+1}}\Big) \\ \iff 1 + &\frac{(w'_k)^{\eta+1}}{d'_{k+1}}\Big(1 - \frac{r_1}{r_k}\Big) < r_1\Big(\frac{d''_{k+1}}{d'_{k+1}}\Big) \end{aligned} \tag{A17}$$

Notice that because of our ordering of firms, $r_1 < r_k$. Thus the left-hand side of (A17) is greater than 1. As a result, we have $1 < r_1\Big(\frac{d''_{k+1}}{d'_{k+1}}\Big)$, thus $\frac{d'_{k+1}}{d''_{k+1}} < r_1$. Applying this result sequentially, we must have $\frac{d'_K}{d''_K} < r_1$. Since $\frac{d'_K}{d''_K} = r_K$, by construction, we have $r_K < r_1$. This contradicts with our ordering of firms, and thus clearly cannot hold. By contradiction, the vector of equilibrium shares and the associated vectors of wages and labor choices must be unique. This contradiction establishes the uniqueness result. (Q.E.D)

Second, we show that a unique equilibrium K^* exists. Suppose there exists two values K_1 and K_2 ($K_1 < K_2$) in this environment that satisfy the equilibrium selection rule. By equilibrium definition we have:

$$\Pi^{K_1}(z^{(K_1)}) \geq 0 > \Pi^{K_1+1}(z^{(K_1+1)}) \quad \text{and} \quad \Pi^{K_2}(z^{(K_2)}) \geq 0 > \Pi^{K_2+1}(z^{(K_2+1)}) \tag{A18}$$

Nonetheless, by productivity ranking, we must have: $z^{(K_1+1)} \geq z^{(K_2)}$ since $K_1 < K_2$. Combine with the profit monotonicity condition proved in section A.2, we must have:

$$0 > \Pi^{K_1+1}(z^{(K_1+1)}) \geq \Pi^{K_2}(z^{(K_2)}) \geq 0. \tag{A19}$$

This condition clearly cannot hold. Therefore the equilibrium K^* is unique. From (A4), and given the set of firm productivies, K^* determines all firm-level variables in equilibrium. (Q.E.D)

A.4 Proposition 2

Proposition 2 states three main results: ${K^*}'(\tau^O) \geq 0$, and for $z \geq z^{(K^*)}$, $s'(\tau^O, .) \leq 0$, $\tilde{\chi}^{e'}(\tau^O, .) \leq 0$.

We first show that ${K^*}'(\tau^O) \geq 0$. To begin with, notice that holding the number of firms K^* fixed, the vector of equilibrium shares is the unique solution to system of equations (A11) and thus, does not change in response to change in aggregate conditions i.e. change in $\mathbf{\Lambda}$ in equation (A4).

Recall from equation (20) that $\mathbf{P}'(\tau^O) \geq 0$, because lower output tariff increases competition in product market and decreases the aggregate price. This leads to lower $\mathbf{\Lambda}$. Plugging this condition into equation (A4), it is straightforward to observe that equilibrium wages decrease and that equilibrium profits decrease along with the wage conditioning on the same shares. Thus, $\Pi'(\tau^O, .) \geq 0$ for all z, holding the number of firms fixed. We now change K^* and show that ${K^*}'(\tau^O) \geq 0$ by contradiction.

Suppose there exists two alternative scenarios of trade policy environment $\tau_1^O > \tau_2^O$, such that $K^*(\tau_1^O) \equiv K_1^* < K^*(\tau_2^O) \equiv K_2^*$. Consider the firm $z^{(K_1^*+1)}$. By the equilibrium definition in proposition 1, we must have $\Pi^{K_1^*, \tau_1^O}(z^{(K_1+1)}) < 0$ and $\Pi^{K_2^*, \tau_2^O}(z^{(K_1+1)}) \geq 0$. Noticing also that from the profit monotonicity condition, we must have $\Pi^{K_2^*}(z^{(K_1+1)}) \leq \Pi^{K_1^*}(z^{(K_1+1)})$. Combining these inequalities, we have:

$$\Pi^{K_1^*, \tau_1^O}(z^{(K_1+1)}) < 0 \leq \Pi^{K_2^*, \tau_2^O}(z^{(K_1+1)}) \leq \Pi^{K_1^*, \tau_2^O}(z^{(K_1+1)}) \leq \Pi^{K_1^*, \tau_1^O}(z^{(K_1+1)}). \tag{A20}$$

Expressions in (A20) clearly can not hold. Therefore, it must be true that as $\tau_1^O > \tau_2^O$, then $K_1^* \geq K_2^*$. (Q.E.D)

Second, we show that market shares and distortions increase for all incumbent firms as τ^O decreases. Consider a reduction in τ^O, and suppose that the reduction is large enough to induce exits of at least one firm $z^{(K*)}$. We will show that $s^{K^*}(z^{(k)}) \geq s^{K^*-1}(z^{(k)})$ for all $k \leq K^* - 1$. It is cumbersome to show this result analytically using calculus. However, we could utilize the proof for uniqueness in section A.3.

Again, in subsequent part of this section, productivity ranking of firms will not matter for the proof, thus I index firms by subscript k rather than productivity z. Let's the vector of equilibrium shares when $K = K^*$ be $(s'_1, ..., s'_{K^*-1}, s'_{K^*})$ and the vector of equilibrium shares when $K = K^* - 1$ be $(s''_1, ..., s''_{K^*-1})$. Here, $s'_{K^*} = s^{K^*}(z^{(K*)})$. We will show by contradiction that $s'_k \leq s''_k$ for all $1 \leq k \leq K^* - 1$. Suppose otherwise that $s'_\kappa > s''_\kappa$ for some $1 \leq \kappa \leq K^* - 1$.

For any $k = 1, \ldots, K^* - 1$, let $r_k = \left(\frac{m(s'_k)\tilde{z}_k}{m(s''_k)\tilde{z}_k}\right)^{\eta+1}$ and reshuffle firms by index k such that

$r_k \leq r_l$ for any $k \leq l$.[52] Similar to the uniqueness proof, we must have $r_\kappa < 1$ and thus $r_1 < 1$, since $m(.)$ is a decreasing function. For any k, denote:

$$d'_k \equiv (m(s'_k)\tilde{z}_k)^{\eta+1} + \ldots + (m(s'_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1} + (m(s'_{K^*})\tilde{z}_{K^*})^{\eta+1} \tag{A21}$$
$$d''_k \equiv (m(s''_k)\tilde{z}_k)^{\eta+1} + \ldots + (m(s''_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1}$$

Following similar steps as in the the uniqueness proof, we can show that $\frac{d'_k}{d''_k} < r_1 \ \forall\ 2 \leq k \leq K^* - 1$. When $k = K^* - 1$, we must have:

$$r_{K^*-1} = \frac{(m(s'_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1}}{(m(s''_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1}} \leq \frac{(m(s'_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1} + (m(s'_{K^*})\tilde{z}_{K^*})^{\eta+1}}{(m(s''_{K^*-1})\tilde{z}_{K^*-1})^{\eta+1}} \tag{A22}$$
$$= \frac{d'_{K^*-1}}{d''_{K^*-1}} < r_1$$

Inequalities in (A22) clearly cannot hold due to our reshuffling. By contradiction, we must have $s'_k \leq s''_k$ for all $1 \leq k \leq K^* - 1$ i.e. labor market shares of incumbent firms increase as some firms exit the market due to lower tariff τ^O. From equation (14), since distortion is an increasing function of share, we also have labor market distortions increase. These results conclude the proof for proposition 2. (Q.E.D)

A.5 Proposition 3

The proof for proposition 3 follows straightforwardly from the proof for proposition 2. The only difference now is that the impact of input tariff is magnified by $\frac{1-\alpha}{\alpha}$, the relative factor shares between labor and intermediate input.

To see this, notice that change in input tariff τ^I affects the competition in the labor market through a similar channel as output tariff τ^O, that is they both affect the labor demand $MRPL$. Therefore, all the arguments in the proof for proposition 2 apply. The goal is now to show that the effect of input tariff on the competition in the labor market could be much larger than that of output tariff, especially if the production uses intermediate input heavily. From the production function in equation (22), I can derive the modified $MRPL$ of firm z as:

$$MRPL(z) = z^{\frac{\gamma-1}{\gamma}} L(z)^{\alpha(\frac{\gamma-1}{\gamma})-1} M(z)^{(1-\alpha)(\frac{\gamma-1}{\gamma})} \alpha(\frac{\gamma-1}{\gamma}) \mathbf{P}^{\frac{\gamma-1}{\gamma}} \mathbf{I}^{\frac{1}{\gamma}} \tag{A23}$$

From equation (A23), conditional on the same amount of labor $L(z)$, change in labor demand

[52] Here, $\tilde{z}_k = z_k^{\frac{\gamma-1}{\gamma+\theta}}$

is determined by:

$$d\log(MRPL(z)) = (1-\alpha)(\frac{\gamma-1}{\gamma})d\log(M(z)). \tag{A24}$$

From the FOC condition of firm z with respect to $M(z)$, we obtain:

$$d\log(M(z)) = -\frac{1}{1-(1-\alpha)(\frac{\gamma-1}{\gamma})}d\log(1+\tau^I). \tag{A25}$$

Combine (A24) and (A25), change in labor demand as a result of change in input tariff is:

$$d\log(MRPL(z)) = -\frac{(1-\alpha)(\frac{\gamma-1}{\gamma})}{1-(1-\alpha)(\frac{\gamma-1}{\gamma})}d\log(1+\tau^I). \tag{A26}$$

Note that in the above equation, the magnified factor $\frac{(1-\alpha)(\frac{\gamma-1}{\gamma})}{1-(1-\alpha)(\frac{\gamma-1}{\gamma})} \to \frac{1-\alpha}{\alpha}$ as $\gamma \to \infty$ i.e. perfect competition in product market. Equation (A26) conludes the proof for proposition 3. (Q.E.D)

B. Production Function Estimation

This section provides supplementary notes for production function estimation procedure in section 3. In particular I provide the detailed derivations of firm's profit maximization problem, timing assumptions as well as the moment conditions for the second GMM stage. A firm maximizes its profit with respect to material input conditional on its information set in period t, denoted by $\mathbb{I}_t$ as follow:

$$\max_{M_t} \mathbb{E}[F(k_t, l_t, m_t) * e^{(\omega_t+\varepsilon_t)}|\mathbb{I}_t] - p_{M_t}M_t, \tag{B1}$$

where $F(.)$ and M_t are the exponential counterparts of $f(.)$ and m_t in equation (26) respectively. p_{M_t} is the price of material, taken as given by the firm. Taking FOC of this problem yields:

$$\frac{\partial}{\partial M_t}F(k_t, l_t, m_t)e^{\omega_t}E[e^{\varepsilon_t}] - p_{M_t} = 0. \tag{B2}$$

Taking the log version of the above equation, I obtain:

$$\begin{aligned} log(s_t^M) \equiv log\frac{p_{M_t}M_t}{R_t} &= logE[e^{\varepsilon_t}] + log\frac{\partial}{\partial m_t}f(k_t, l_t, m_t) - \varepsilon_t \\ &= log\frac{\partial}{\partial m_t}f(k_t, l_t, m_t) - \varepsilon_t. \end{aligned} \tag{B3}$$

The second equality of (B3) (also equation (27)) follows under the assumption that $E[e^{\varepsilon_t}] = 1$ or $E[\varepsilon_t] = 0$. In my empirical implementation, I estimate $\hat{\varepsilon}_t$ and correct for any asymmetry in the measurement error ε_t (see also in Gandhi, Navarro and Rivers (2020)). In the Chinese firm-level data, the estimated $\hat{\varepsilon}_t$ exhibits little asymmetry and requires minimum correction, i.e. $E[e^{\hat{\varepsilon}_t}] \approx 1$ for most industries.

In what follows, $\mathbf{v}_t$ is a vector of additional state variables that I control for including year, location, industry, firm's ownership, export status and tariff levels. The timing assumptions of the GNR productivity model is as follows:

- At the end of period $(t-1)$, the firm chooses $(k_t, l_t, \mathbf{v}_t)$ and whether to exit at t.
- At the beginning of period t, η_t (and hence ω_t) realizes. The firm observes their productivity for period t.
- The firm optimally chooses m_t, after which ε_t realizes and completely determines r_t.
- At the end of t, the firm chooses $(k_{t+1}, l_{t+1}, \mathbf{v}_{t+1})$ and whether to exit at $t+1$, repeating the same process.

The moment conditions for the second GMM stage are:

$$E\left[\eta_t \otimes \left(\begin{bmatrix} 1 \\ \Psi_{t-1} \\ C_t(.) \\ C_{t-1}(.) \end{bmatrix} \right)\right] = 0 \tag{B4}$$

C. Data Construction and Filtering

The China's Annual Survey of Industrial Enterprises (ASIE) data record firms' balance sheets information and contain a firm-specific identifier (ID). Firms could change ID over time due to various reasons (e.g. due to M&A activity). I match firms over the years in the sample first based on their IDs. After matching on IDs, I match firms based on name, zip code, telephone number, and legal person representatives concurrently. The matching code follows the published code in Brandt, Biesebroeck and Zhang (2014).

After matching, my cleaning procedure is performed as follow, in sequential steps:

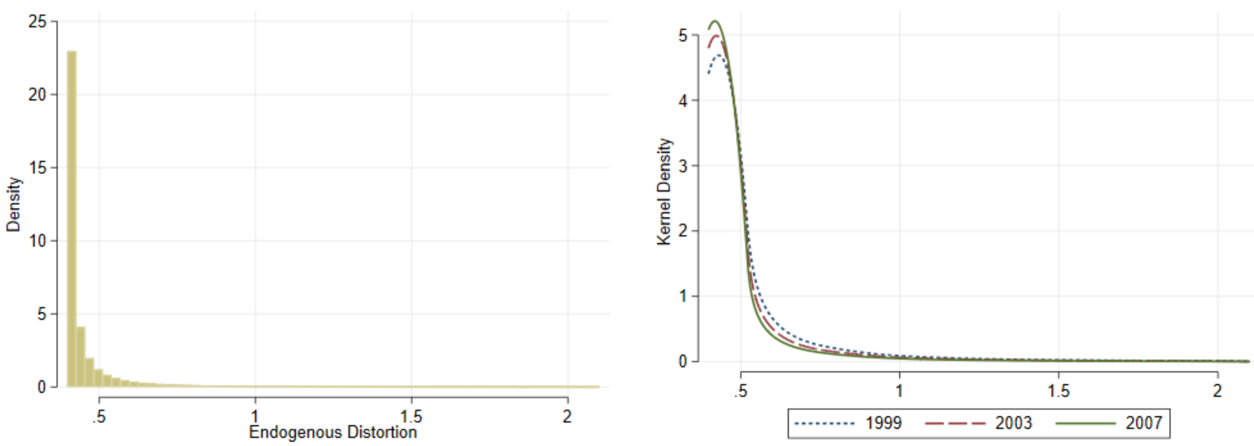
- Step 1: Drop all firms with missing values of key variables: output (revenue), real capital stocks, employment, materials, wage-bill, export status.

- Step 2: Drop all firms with values of key variables outside of range $0.5th$ and $99.5th$ percentiles. These variables include: output (revenue), real capital stocks, employment, materials, wage-bill, average wage, labor share and input share.
- Step 3: Drop all firms with labor share and input share outside of range $[0, 1]$.
- Step 4: Drop all firms with output (revenue) below 5 million Renminbi (RMB) and employment below 8 workers to keep consistent size thresholds.
- Step 5: Drop all firms that switch 2-digit industries. This ensures a consistent production function estimation at 2-digit industry level and correct tariff exposures for each firm.

My cleaning procedures are similar to standard practices in the literature that uses ASIE data, see for example Brandt, Biesebroeck and Zhang (2014) and Brandt et al. (2017).

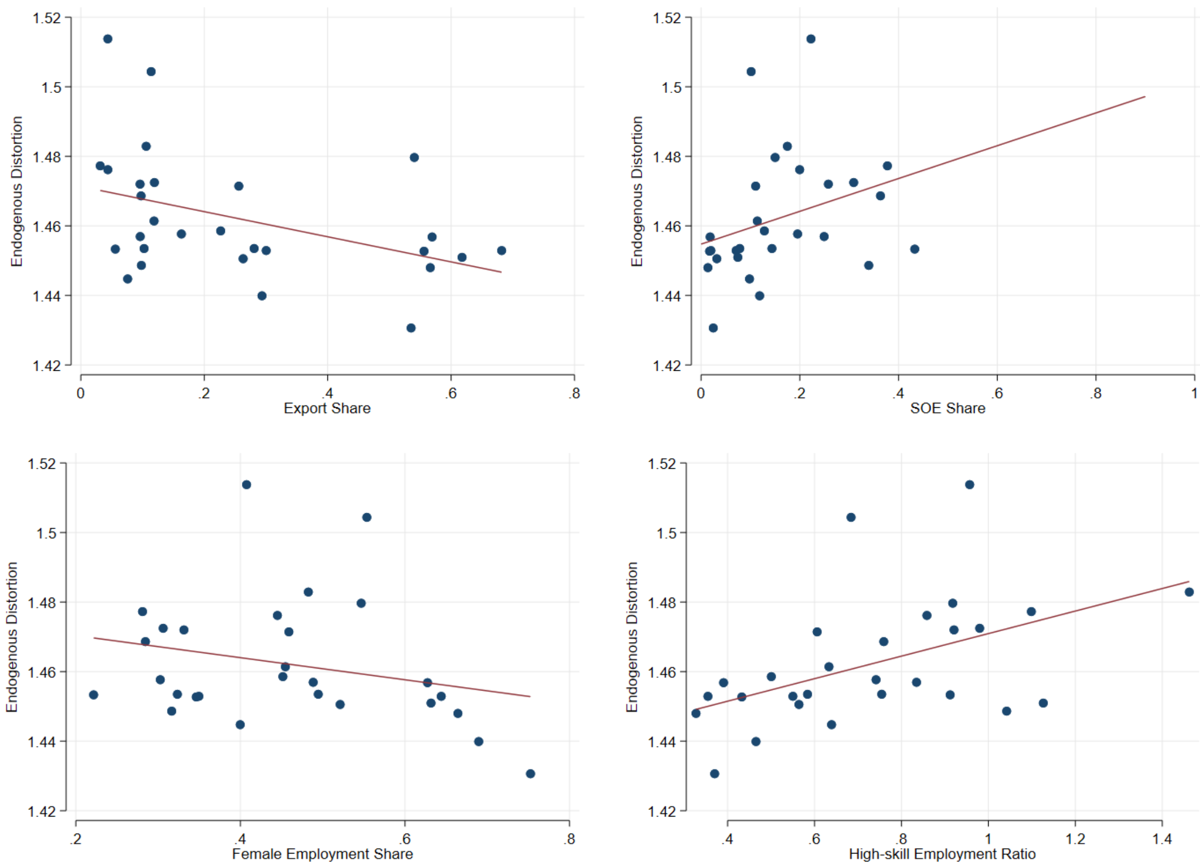
D. Additional Figures and Tables

Figure D1: Distribution of the Endogenous Distortion ($log(\tilde{\chi}^e(s_{ijlt}))$)



Note: The figure is a counterpart of Figure 2, but for the endogenous distortion ($\tilde{\chi}^e(s_{ijlt})$) rather than the overall distortion ($\tilde{\chi}_i$). The left panel shows the distribution of distortion across all firm-year observations. The right panel displays the evolution of distortion distribution over three equidistant years: 1999, 2003, 2007.

Figure D2: Endogenous Distortion ($\tilde{\chi}^e(s_{ijlt})$) and Industry Characteristics (in year 2004)



Note: The figure is a counterpart of Figure 3, but for the endogenous distortion ($\tilde{\chi}^e(s_{ijlt})$) rather than the overall distortion ($\tilde{\chi}_i$). It shows the correlations between the measured endogenous distortion ($\tilde{\chi}^e(s_{ijlt})$) and (2-digit) industry characteristics in 2004. The industry characteristics include: export share, state ownership (SOE) share, female employment share and high-skill employment ratio. High-skill employment ratio is defined as the ratio between high-school and secondary-school degree workers. The figure is based on the data in 2004, because this is the only year that the employment composition information is available.

Table D1: Aggregate Summary Statistics of the Chinese Firm-level Data

Year	No. of Firms	VA	Sales	Output	Employment	Export	Fixed Assets (Net)
1998	165118	1.94	6.54	6.77	56.44	1.08	4.41
1999	162033	2.16	7.06	7.27	58.05	1.15	4.73
2000	162882	2.54	8.37	8.57	53.68	1.46	5.18
2001	171256	2.83	9.19	9.54	54.41	1.62	5.54
2002	181557	3.30	10.86	11.08	55.21	2.01	5.95
2003	196220	4.20	13.95	14.23	57.48	2.69	6.61
2004	279092	6.62	19.78	20.17	66.22	4.05	7.97
2005	271835	7.22	24.69	25.16	69.31	4.77	8.95
2006	301961	9.11	31.08	31.66	73.49	6.05	10.58
2007	336768	11.70	39.76	40.51	78.75	7.34	12.34

Note: The table reports the aggregate summary statistics of the Chinese firm-level data, prior to cleaning procedures. Employment is in millions of workers. All monetary values are denoted in trillions Renminbi (RMB).

Table D2: Revenue Elasticities and Labor Market Distortion by Industry (ACF-Translog)

Industry	Capital	Labor	Material	$\tilde{\chi}_i$ (**Mean**)
13. Food Processing	0.19	0.19	.	6.94
14. Food Production	0.21	0.19	.	3.58
15. Beverage	0.21	0.06	.	1.23
16. Tobacco	0.39	0.33	.	3.10
17. Textile	0.22	0.21	.	3.50
18. Garments	0.14	0.35	.	3.16
19. Leather	0.17	0.26	.	2.86
20. Timber	0.25	0.25	.	4.68
21. Furniture	0.14	0.34	.	4.73
22. Paper-making	0.23	0.21	.	4.24
23. Printing	0.19	0.08	.	1.03
24. Cultural	0.16	0.26	.	2.80
25. Petroleum Processing	0.25	0.19	.	4.75
26. Raw Chemical	0.23	0.13	.	2.71
27. Medical	0.27	0.15	.	2.60
28. Chemical Fibre	0.33	0.19	.	5.22
29. Rubber	0.19	0.10	.	1.55
30. Plastic	0.20	0.22	.	3.81
31. Nonmetal Products	0.19	0.04	.	0.47
32. Processing of Ferrous	0.30	0.26	.	7.13
33. Processing of Nonferrous	0.22	0.20	.	5.79
34. Metal Products	0.21	0.19	.	2.93
35. Ordinary Machinery	0.19	0.12	.	1.72
36. Special Equipment	0.16	0.14	.	1.87
37. Transport Equipment	0.25	0.27	.	4.32
39. Electric Machinery	0.25	0.23	.	4.05
40. Electronic and Telecom	0.21	0.35	.	4.83
41. Measuring Instruments	0.13	0.13	.	1.28
42. Art Work	0.15	0.17	.	1.73
All Industry	0.21	0.20	.	3.40

Note: This table reports estimates of the revenue elasticities of factor inputs: capital and labor, and the measured *overall distortion* ($\tilde{\chi}_i$), using the translog production function estimation procedure in Ackerberg, Caves and Frazer (2015) (ACF). All statistics are the mean of respective distributions. The ACF method assumes that production function is Leontief, i.e. perfect complementarity between material and other factors. The table trims observations above and below the 1^{st} and 99^{th} percentiles.

Table D3: China's Tariffs Evolution from 1998-2007

Year	Output Tariff (τ^O)		Input Tariff (τ^I)	
	Average	Std. Deviation	Average	Std. Deviation
	(1)	(2)	(3)	(4)
1998	0.16	0.09	0.11	0.03
1999	0.16	0.09	0.11	0.03
2000	0.16	0.09	0.11	0.03
2001 (WTO)	0.15	0.08	0.10	0.03
2002	0.12	0.07	0.08	0.03
2003	0.11	0.07	0.07	0.02
2004	0.10	0.06	0.06	0.02
2005	0.09	0.06	0.06	0.02
2006	0.09	0.06	0.06	0.02
2007	0.09	0.06	0.06	0.02
Total	0.12	0.08	0.08	0.03

Note: All tariffs are computed as the natural log of 1 plus the ad valorem tariffs i.e. $ln(1+\tau)$. Input tariffs are computed as weighted averages of output tariffs, using input shares from China's Input-Output table in 2002 as weights.